

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 3, 2006, 00:41:37 ; Search time 941 Seconds
(without alignments)
2899.994 Million cell updates/sec

Title: US-09-390-846-2

Perfect score: 1688

Sequence: 1 MAVFKNTRPKIAMVSGMI.....GSIDEVKEMQKAIALDASK 330

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/abss/ABSSWEB.spool/US09390846/runat_02032006_104243_5094/app_query.fasta_1
-DB=Published Applications NA Main -QFWT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs03h
-USER=US09390846 @CGN 1.1 1026 @runat_02032006_104243_5094 -NCPU=6 -ICPU=3
-NO.WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3:	/cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
4:	/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5:	/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6:	/cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7:	/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8:	/cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9:	/cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10:	/cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1034.5	61.3	1785	3	US-09-216-393-311
2	1034.5	61.3	1785	3	US-09-216-393-311
3	1034.5	61.3	1785	6	US-10-321-856-311
4	1034.5	61.3	1785	6	US-10-321-856-311
5	880.5	52.2	960	6	US-10-369-493-41525
6	873.5	51.7	957	6	US-10-369-493-35128
7	873.5	51.7	960	6	US-10-369-493-38176

Alignment Scores:

Pred. No.:	4.91e-111	Length:	1785
Score:	1034.50	Matches:	193
Percent Similarity:	77.3%	Conservative:	55
Best Local Similarity:	60.1%	Mismatches:	72
Query Match:	61.3%	Indels:	1
DB:	3	Gaps:	1

ALIGNMENTS

RESULT 1

US-09-216-393-311
; Sequence 311, Application US/09216393
; Patent No. US20010014447A1
; GENERAL INFORMATION:
; APPLICANT: Milhausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216.393
; CURRENT FILING DATE: 1998-12-18
; EARLIER APPLICATION NUMBER: 08/994.825
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 311
; LENGTH: 1785
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(75)
US-09-216-393-311

8	873.5	51.7	960	6	US-10-369-493-38561	Sequence 38561, A
9	862.5	51.1	933	6	US-10-369-493-44435	Sequence 44435, A
10	862	51.1	969	6	US-10-369-493-35747	Sequence 35747, A
11	850.5	50.4	963	6	US-10-369-493-40766	Sequence 40766, A
12	848.5	50.3	960	6	US-10-369-493-31700	Sequence 31700, A
13	847.5	50.2	960	6	US-10-369-493-34210	Sequence 34210, A
14	740	43.8	939	6	US-10-369-493-46924	Sequence 46924, A
15	709	42.0	945	6	US-10-369-493-41082	Sequence 41082, A
16	697	41.3	936	6	US-10-369-493-40205	Sequence 40205, A
17	689	40.8	855	6	US-10-369-493-42814	Sequence 42814, A
18	675.5	40.0	959	6	US-10-369-493-37272	Sequence 37272, A
19	672	39.8	927	6	US-10-369-493-32732	Sequence 32732, A
20	652.5	38.7	912	6	US-10-369-493-34088	Sequence 34088, A
21	649.5	38.5	978	6	US-10-369-493-41907	Sequence 41907, A
22	641	38.0	978	7	US-10-724-972A-3485	Sequence 3485, Ap
23	635	37.6	912	6	US-10-369-493-34859	Sequence 34859, A
24	635	37.6	951	6	US-10-369-493-43537	Sequence 43537, A
25	634	37.6	975	6	US-10-369-493-26509	Sequence 26509, A
26	628	37.2	897	6	US-10-369-493-33451	Sequence 33451, A
27	626.5	37.1	915	6	US-10-369-493-34654	Sequence 34654, A
28	625.5	37.1	927	6	US-10-369-493-42545	Sequence 42545, A
29	621	36.8	822	3	US-09-974-300-56	Sequence 56, Appl
30	605.5	35.9	1008	6	US-10-369-493-23825	Sequence 23825, A
31	569	33.7	1005	6	US-10-369-493-23812	Sequence 23812, A
32	560.5	33.2	942	6	US-10-369-493-46543	Sequence 46543, A
33	518	30.7	595	6	US-10-369-493-38013	Sequence 38013, A
34	498.5	29.5	678	3	US-09-974-300-4543	Sequence 4543, Ap
c 35	481.5	28.5	1163020	7	US-10-398-221-10	Sequence 10, Appl
c 36	478	28.3	945	6	US-10-369-493-42017	Sequence 42017, A
c 37	473.5	28.1	3011208	7	US-10-398-221-2058	Sequence 2058, Ap
38	462	27.4	936	6	US-10-369-493-40366	Sequence 40366, A
39	460	27.3	966	6	US-10-369-493-46760	Sequence 46760, A
40	458	27.1	936	6	US-10-369-493-40390	Sequence 40390, A
41	455	27.0	957	6	US-10-154-460-45	Sequence 45, Appl
42	440.5	26.1	1062	3	US-09-938-842A-1938	Sequence 1938, Ap
43	440.5	26.1	1062	3	US-09-938-842A-1938	Sequence 1938, Ap
44	439.5	26.0	960	6	US-10-369-493-26756	Sequence 26756, A
45	439	26.0	936	6	US-10-369-493-40165	Sequence 40165, A

US-09-390-846-2 (1-330) x US-09-216-393-311 (1-1785)

Qy 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
Db 109 AGAAGAAGTGGCCATGTTGGCTCTGGCATGATTGGTGGCATTATGGCTTACTGTGC 168

Qy 29 SerLeuArgGluLeuGlyAspValValLeuPheAspValProAsnMetProMetGly 48
Db 169 GCTCTCGTGAGCTCGCTGAGCTGTTCTCTACGATGTTGTCAAAGGATGCGCGAGG 228

Qy 49 LysAlaMetAspIleSerHisAsnSerValValAspThrGlyIleThrValTyrGly 68
Db 229 AAGGCTCTTACGACCGATGTGACCTCGGTGCGATCGATCAACCAAGCTTCCGTCG 288

Qy 69 SerAsnSerTyrGlu---CysLeuLysGlyAlaAspValValIleThrAlaGlyIle 87
Db 289 GAGTACTCTTACGAGCGCGCTCACCGGTGCGGACTGCGTTATCGTTACCGCCG 348

Qy 88 ThrLysIleProGlyLysSerAspLysGluTTPSerArgMetAspLeuProValAsn 107
Db 349 ACCAAGTGCAGGCGGCAAGCCGACTCCGAGTGGAGCCGAAACGATGCTCCCGTTCA 408

Qy 108 IleLysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAlaPheVal 127
Db 409 TCGAAGATCATTCGCGAGATCGGTGAGATCAAGAACTACTGCCCCAAGACCTTCATC 468

Qy 128 IleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeu 147
Db 469 ATCGTGTGTACCAACCGCTGCGATCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 528

Qy 148 ProHisIleArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMet 167
Db 529 CCAGCAACATGATCTGCGGTATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 588

Qy 168 IleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHis 187
Db 589 GTCCGCGACGCGTGTCTGCTCTCTCCCGCGAGCTCCAGGCCACCGCTATCGGCAC 648

Qy 188 GlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGlu 207
Db 649 GCGGACTGTCATGCTCCCGCTTGTCTCCGATCAATTTACCGTGAACGACTTACCGCA 708

Qy 208 PheValLysGlyTyrPheLysGlnGluValAspAlaValGlnLysThrLys 227
Db 709 TTATCAAGGACCGCGTAGTGCAGGAGACGAGCTCGAGGAGTCTGCTGAGCACACCAA 768

Qy 228 ValAlaGlyGlyIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaProGly 247
Db 769 GTGCTCGCGCGAGATCGTCCGCTTCTCCGCGAGGTTCCGCTTACTTACGCCCCCGCC 828

Qy 248 AlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCys 267
Db 829 GCATCCGCTCTGCGCATGCGCAACATCTCTTGAACGAGCAAAAGGCGCTATCCCGTGC 888

Qy 268 SerCysTyrLeuGlnGlyIntTyrGlyValGlnAsnHisTyrLeuGlyValProCysVal 287
Db 889 AGTGTGTCTGCAACGAGAGTACGCTTGAAGACATGTTCAATGCTTCCCGCGCTC 948

Qy 288 IleGlyGlyArgGlyValGluLysIleIleGluLeuLeuThrAlaGlnGluArgGln 307
Db 949 ATTGGAGCGCGCGCATCGAGCGCTCATCGAGCTCGAGCTGAACGAGGAGAGAAG 1008

Qy 308 GluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaLeuAsp 327
Db 1009 CAGTTCAGAAAGTCCGTCGACGCGTATCGGCGCTCAACAAAGCGGCTTCTGCTCTTCA 1068

Qy 328 Ala 328
Db 1069 GCG 1071

RESULT 2
US-09-216-393-311

QY 228 ValAlaGlyGlyGluValValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaProGly 247
 DB 769 GTGTCGGCGGCGAGATCGTCCTCCCGCCAGGGTTCGCTTACTACGCCCGCCGCC 828
 QY 248 AlaSerAlaIleGlnMetAlaGluSerTyrLeuLeuAspArgValMetValCys 267
 DB 829 GCATCGGCTGTCCGATGCAACATCTTCTTGAACGACGAAAGCGCTCATCCCGTGC 888
 QY 268 SerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysVal 287
 DB 899 AGTGTTACTGCAACGAGGAGTACGGCTTGAAGGACATGTTCAATGGTCTCCCGCGCTC 948
 QY 288 IleGlyGlyArgGlyValGluValIleGluLeuLeuThrAlaGlnGluArgGln 307
 DB 949 ATTGAGGCGCGGATCGATCGCGCTCATCGAGCTCGAGCTGAACGAGGAGGAGAG 1008
 QY 308 GluLeuGlnGlySerIleAspGluValValGluMetGlnLysAlaIleAlaLeuAsp 327
 DB 1009 CAGTTCCAGAGTCCGTCGACGAGTTCATGGCGCTCAACAGGCGGTGCTCTTCAG 1068
 QY 328 Ala 328
 DB 1069 GCG 1071

RESULT 3
 US-10-321-856-311
 ; Sequence 311, Application US/10321856
 ; Publication No. US20030194393A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Milhausen, Michael James
 ; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF
 ; FILE REFERENCE: TX-1-C2-1
 ; CURRENT APPLICATION NUMBER: US/10/321,856
 ; PRIOR FILING DATE: 2002-12-17
 ; PRIOR APPLICATION NUMBER: 09/216,393
 ; PRIOR FILING DATE: 1998-12-18
 ; PRIOR APPLICATION NUMBER: 08/994,825
 ; PRIOR FILING DATE: 1997-12-19
 ; NUMBER OF SEQ ID NOS: 366
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 311
 ; LENGTH: 1785
 ; TYPE: DNA
 ; ORGANISM: Toxoplasma gondii
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(75)
 ; OTHER INFORMATION:
 US-10-321-856-311

Alignment Scores:
 Pred. No.: 4,91e-111 Length: 1785
 Score: 1034.50 Matches: 193
 Percent Similarity: 77.3% Conservative: 55
 Best Local Similarity: 60.1% Mismatches: 72
 Query Match: 61.3% Indels: 1
 DB: 6 Gaps: 1

US-09-390-846-2 (1-330) x US-10-321-856-311 (1-1785)

QY 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
 DB 109 AGAAAGAGGTGGCCATGATGGCTCTGGCATGATTGGTGGCATTATGGCTACCTGTGC 168
 QY 29 SerLeuArgGluLeuGlyAspValValLeuPheAspValProAsnMetProMetGly 48
 DB 169 GCTCTCCGTGAGCTCGCTGACGCTGCTTCTACGATGTGTCAAGGATGATGCCGAGGT 228
 QY 49 LysAlaMetAspIleSerHisAsnSerValValAspThrGlyIleThrValTyrGly 68
 DB 229 AAGGCTCTTGACCTGAGCCATGTGACCTCGGTGCTGACACCAAGCTTTCGTCGCTCT 288
 QY 69 SerAsnSerTyrGlu---CysLeuLysGlyAlaAspValValIleIleThrAlaGlyIle 87

DB 289 GAGTACTCTTACGAGCGCGCTCACCGGTGCGGACTCGGTTCATCGTTACCCCGCGTCTG 348
 QY 88 ThrLysIleProGlyLysSerAspLysGluTyrPsrArgMetAspLeuLeuProValAsn 107
 DB 349 ACCAAGGTGCGCGGCAAGCCGACCTCCGAGTGGAGCCGAAACGATCTGCTCCCGTTCAC 408
 QY 108 IleLysIleMetArgGluValGlyAlaIleLysSerTyrCysProAsnAlaPheVal 127
 DB 409 TCGAAGATCATTCGCGGAGATCGGTGAGAACATCAAGAAAGTACTGCCCAAGACCTTCATC 468
 QY 128 IleAsnIleThrAspProLeuAspValMetValAlaIleLeuGlnGluSerSerGlyLeu 147
 DB 469 ATCTGGTGACCAACCCCGTGCATGTCATGGTCAAGGTTCATGTGCGGCGCTTCGGCGTC 528
 QY 148 ProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgArgMet 167
 DB 529 CGGACCAACATCATCTCGGGTATGGCTGTCATCTCGACTCTGGTCTGCTCCGCGATAC 588
 QY 168 IleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHis 187
 DB 589 GTGCGCGAGCGGCTGTCTGTCTCTCCCGCGAGCTCCAGGCCACCGTCATCGGCACACAC 648
 QY 188 GlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlnIleProLeuSerGlu 207
 DB 649 GCGGACTGCATGCTCCCGCTTGTCCGGTACATTACCGTGAACGACTACCCGATCCAGAAG 708
 QY 208 PheValLysLysGlyTyrIleLysGlnGluValAspAspIleValGlnLysThrLys 227
 DB 709 TTCATCAGGACGGCGTAGTCACGAGAAAGCAGCTCGAGGAGATCGCTGAGCACAACAAA 768
 QY 228 ValAlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaProGly 247
 DB 769 GTGTCGCGCGAGATCGTCGCTTCTCCGCGAGGGTTCGCTTACTACGCCCGCCGCC 828
 QY 248 AlaSerAlaIleGlnMetAlaGluSerTyrLysLysAspArgLysArgValMetValCys 267
 DB 829 GCATCCGCTGCGCCATGGCAACATCTCTTGAACGACGAAAGCGGTCTATCCCGTGC 888
 QY 268 SerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysVal 287
 DB 889 AGTGTTACTGCAACGAGAGTACCGCTTGAAGGACATGTTCAITGGTCTCCCGCGCTC 948
 QY 288 IleGlyGlyArgGlyValGluValIleGluLeuGluLeuThrAlaGlnGluArgGln 307
 DB 949 ATTGAGGCGCGGATCGATCGCGCTCATCGAGCTCGAGCTGAACGAGGAGGAGAGAG 1008
 QY 308 GluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaLeuAsp 327
 DB 1009 CAGTTCCAGAGTCCGTCGACGAGCTCATGGCGCTCAACAGGCGGTGCTCTTCAG 1068
 QY 328 Ala 328
 DB 1069 GCG 1071

RESULT 4
 US-10-321-856-313
 ; Sequence 313, Application US/10321856
 ; Publication No. US20030194393A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Milhausen, Michael James
 ; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF
 ; FILE REFERENCE: TX-1-C2-1
 ; CURRENT APPLICATION NUMBER: US/10/321,856
 ; PRIOR FILING DATE: 2002-12-17
 ; PRIOR APPLICATION NUMBER: 09/216,393
 ; PRIOR FILING DATE: 1998-12-18
 ; PRIOR APPLICATION NUMBER: 08/994,825
 ; PRIOR FILING DATE: 1997-12-19
 ; NUMBER OF SEQ ID NOS: 366
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 313
 ; LENGTH: 1785

```

; TYPE: DNA
; ORGANISM: Toxoplasma gondii
US-10-856-313

Alignment Scores:
Pred. No.: 4.91e-111 Length: 1785
Score: 1034.50 Matches: 193
Percent Similarity: 77.3% Conservative: 55
Best Local Similarity: 60.1% Mismatches: 72
Query Match: 61.3% Indels: 1
DB: 6 Gaps: 1

US-09-390-846-2 (1-330) x US-10-321-856-313 (1-1785)

Qy 9 ArgProlylleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
Db 109 AGAAGAAGGTGGCCATGATGGCTTGGCATGATGGTGGCACTATGGGTCTACCTGTGC 168

Qy 29 SerLeuArgGluLeuGlyAspValValPheAspValProAsnMetProMetGly 48
Db 169 GCTCTCCGTGAGCTGCGTACGCTGCTTCTTCTACGATGTTGTCAAGAGGTATGCCCGAGGT 228

Qy 49 LysAlaMetAspIleSerHisAsnSerValValAspThrGlyIleThrValTyrgly 68
Db 229 AAGGCTCTTGACCTGAGCCATGTGACCTCCGTTGGTTCGACACCAAGTTTCGTTCGTGCT 288

Qy 69 SerAsnSerTyrglu--CysLeuLysGlyAlaAspValValIleIleThrAlaGlyIle 87
Db 289 GAGTACTCTACGAGCGCGCTACCGTGGGACTGCGTATGTTTACCGCGGTCTG 348

Qy 88 ThrIysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuProValAsn 107
Db 349 ACCAAGGTGCGGGGCAAGCCGACTCCGAGTGGAGCCGAAACGATCTGCTCCCGTTCAAC 408

Qy 108 IleIysIleMetArgGluValGlyAlaAlaIleLysSerTyrglyProAsnAlaPheVal 127
Db 409 TCGAAGATCATTCGGGAGATCGGTGAGAACATCAAGAGTACTGCCCCCAAGACCTTCATC 468

Qy 128 IleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeu 147
Db 469 ATCGTGGTCAACCAACCCGCTGGACTGTCATGTCGAGGTCAATGTCGGAGCCCTTGGCGTC 528

Qy 148 ProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMet 167
Db 529 CCGACCAACATGATCGCGGTATGCGCTGCACTGCTGCACTCTGCTGCTTCCGCGATAC 588

Qy 168 IleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHis 187
Db 589 GTCCGCGAGCGCTGCTGCTCTCTCCCGGACGTCGAGGCCACCGTTCATCGGCACACAC 648

Qy 188 GlyAspHisMetValProLeuSerArgTyrglyAlaThrValAsnGlyIleProLeuSerGlu 207
Db 649 GCGCACTGATGTCGCTGTCGCGTACATTCGCTGACGACTACCGTCCAGAG 708

Qy 208 PheValLysLysGlyTrpIleLysGlnGluValAspAspIleValGlnLysThrLys 227
Db 709 TTCATCAAGAGCGGCTAGTACCGAGAGAGCAGCTCGAGGAGATCGCTGAGCACACCAA 768

Qy 228 ValAlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrglyAlaProGly 247
Db 769 GTGCTGCGCGGAGATGCTCGCTTCTTCGCGCCAGGGTTCGCTTACACCCCGCCGCC 828

Qy 248 AlaSerAlaIleGlnMetAlaGluSerTyrglyLeuLysAspArgLysArgValMetValCys 267
Db 829 GCATCCGCTGTCCCATGCCATCCTCTTGAACGACGAAAGCGGTATCCCGTGC 888

Qy 268 SerCysTyrglyGlnGlyGlnTyrglyValGlnAsnHisTyrglyLeuValProCysVal 287
Db 889 AGTGGTGACTGCAACCGGAGTACGCTTGAAGGACATGTTTCAATGGTCTCCCGCGCTC 948

Qy 288 IleGlyGlyArgGlyValGluGlyIleGluLeuGluLeuThrAlaGlnGluArgGln 307
Db 949 ATTGAGGCGCGGCATCGAGCGCGTCTACGAGCTCGAGCTGAACGAGGAGGAGGAAG 1008

```

```

Qy 308 GluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaLeuAsp 327
Db 1009 CAGTTCAGAAAGTCCGTCGACGACGTCATGCGCGCTCAACAGCGGTTCCTCTTCAG 1068

Qy 328 Ala 328
Db 1069 GCG 1071

RESULT 5
US-10-369-493-41525
; Sequence 41525, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 41525
; LENGTH: 960
; TYPE: DNA
; ORGANISM: SPHINGOMONAS
US-10-369-493-41525

Alignment Scores:
Pred. No.: 2.88e-93 Length: 960
Score: 880.50 Matches: 173
Percent Similarity: 70.7% Conservative: 54
Best Local Similarity: 53.9% Mismatches: 89
Query Match: 52.2% Indels: 1
DB: 6 Gaps: 1

US-09-390-846-2 (1-330) x US-10-369-493-41525 (1-960)

Qy 9 ArgProlylleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
Db 7 CGCAAGAAGATCCGCGCTGATCGGTCGCGCATGATCGGTGGCACCTTCGCGACCTTCGCT 66

Qy 29 SerLeuArgGluLeuGlyAspValValPheAspValProAsnMetProMetGly 48
Db 67 GCATCAAGNACTGGCGGATATGCTGTTCCACATTCGCGAGGGCCAGCGCGCGC 126

Qy 49 LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyrgly 68
Db 127 AAGCGCTCGACATCGCACAGTCGCGACCGGTTCGAAGCTTCGATGCCCAACCTCAAGGCG 186

Qy 69 SerAsnSerTyrglyCysLeuLysGlyAlaAspValIleIleThrAlaGlyIleThr 88
Db 187 GCGAACAGTACGAGGACATTCGCGCGCGCGCACGTCGTCATCGTCACCGCGGTATCCCG 246

Qy 89 LysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProValAsnIle 108
Db 247 CGCAAGCGCGGC-----ATGAGCGCGGACGATCTTCTCAAGACCAACCTG 291

Qy 109 LysIleMetArgGluValGlyAlaAlaIleLysSerTyrglyCysProAsnAlaPheValIle 128
Db 292 GCGGTGATGAAGCGGTGGCGGAGGCGCATCGCGCACGCGCGCGCGCGCGCGCGCGCGCG 351

Qy 129 AsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeuPro 148
Db 352 TGCATCAACCAACCGCGCTCGACGCGATGTTGGGCGGCTGCGCGGAGTTCCTCGGCGCTCGCG 411

Qy 149 HisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMetIle 168
Db 149 HisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMetIle 168

```


Db	7	CGCAAAAAAATATGCACCTATTATGGTTCTGGCATATCGCGCGACGCTGGCGCATCTCGCC	66
Qy	29	SerLeuArgGluLeuGlyAspValValLeuPheAspValProAsnMetProMetGly	48
Db	67	AGCCTGAAGAACTATGGCGATATCGTCTCTTCGACATCGCGACGGCATCCGCGAGGC	126
Qy	49	LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyGly	68
Db	127	AAGGCTCTGGATATATGCCACTCGCGCCGGTTGAAGGCTTCAATGCMAAGCTCTCCGCG	186
Qy	69	SerAsnSerTyGluCysLeuGlyGlyAlaAspValValIleIleThrAlaGlyIleThr	88
Db	187	GCITCCGATACCGCGCCATCGAAGCGCAGACGCTCTGCATCGTCAACGAGGTGTGCGC	246
Qy	89	LysIleProGlyIysSerAspIysGluTrpSerArgMetAspLeuLeuProValAsnIle	108
Db	247	CGCAAGCCCGGC-----ATGAGCCGCGATGATCTTCTTGGCATCAACCTC	291
Qy	109	LysIleMetArgGluValGlyAlaAlaIleIysSerTyCysProAsnAlaPheValIle	128
Db	292	AAGGTCATGGAAACAGGTGCGCGCCGCGCATCAAGAAATATGCTCCGAACGCTTCGTGATC	351
Qy	129	AsnIleThrAsnProLeuAspValMetValAlaLeuGlnGluSerSerGlyLeuPro	148
Db	352	TGCATCACCAACCCGCTCGACCGCATGTCTGGCGCTTCAGAGTTCTCCGCGCTGCGC	411
Qy	149	HisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgArgMetIle	168
Db	412	AAGAAACAGGTCTGTCGCGATCGCTGGCGTTCTCGACAGCGACGCTTCGCGCTGTTCC	471
Qy	169	AlaAspIysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHisGly	188
Db	472	GCGCAAGAAATTCACAGTTTCGCTCCAGACGCTCACCGCTTCGTTCTCGCGCGTCATGC	531
Qy	189	AspHisMetValProLeuSerArgTyAlaThrValAsnGlyIleProIleSerGluPhe	208
Db	532	GACACATGTGCGCGCTCGCGCTTATTCACCGTTGCGCGCTTCGCGCTTACCATCTC	591
Qy	209	ValIysIysGlyTyPilleLysGlnGluValAspAspIleValIcIlnIysThrLysVal	228
Db	592	GTCNAGATGGCTGTTGACCGCGAACGCTTTCAGCAGATCATTCAGGCGACCGTGAC	651
Qy	229	AlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyTyAlaProGlyAla	248
Db	652	GGCGGCGCGAAATCGTCGGCTCTTGAAGACCGGCTCGGCTATTACGCGCGCGCGCT	711
Qy	249	SerAlaIleGlnMetAlaGluSerTyIleLysAspArgLysArgValMetValCysSer	268
Db	712	TCGGCTATCGAAATGGCGGAATCTCATCTCAAGGACAGAAGCGGTTCTGCGCGCTGCT	771
Qy	269	CysTyIleuGlnGlyGlnTyGlyValIcIlnAsnHisTyIleLeuGlyValProCysValIle	288
Db	772	GCCACCTTTTCGGGCCAGTATGGCGTGAACATGATATGCGGCGTCCCAACCATCATC	831
Qy	289	GlyGlyArgGlyValGlnLysIleIleGlnLeuGluLeuThrAlaGlnArgGlnGlu	308
Db	832	GGTCCCGCGGTATCGAGCGCGTATCGAGATCGAATCTGAACAGGAGAGAGACCGCC	891
Qy	309	LeuGlnGlySerIleAspGluValLysGluMetGlnLysAla	322
Db	892	TTCCAGAAATCCGTCGGCGCTGTCGCTGGTCTTTTCGGAAGCC	933

RESULT 6
US-10-369-493-35128
; Sequence 35128, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

[illegible][illegible]

895
GAGGACGGAGAGAGAGAGAGAGCGCC

Qy LeuGlnGlySerIleAspGluValLysGluMetGlnLysAla 322
||| ||| |||
832 GGGTGGCGGTATCTGGCGGGCATCGAGATCGATCTGAATCAAGGAGAGAGAGAGCGCC

Db 892 TTCCAGAAATCCGTCGGCGCTGCTGGCTCTTTTGGAGGCC 933

RESULT 7
US-10-369-493-38176
; Sequence 38176, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei

APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.

APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 38176
; LENGTH: 960
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-38176

Alignment Scores:
Pred. No.: 1,93e-92 Length: 960
Score: 873.50 Matches: 167
Percent Similarity: 72.3% Conservative: 60
Best Local Similarity: 53.2% Mismatches: 82
Query Match: 51.7% Indels: 5
DB: 6 Gaps: 1

US-09-390-846-2 (1-330) x US-10-369-493-38176 (1-960)

QY 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyGlyThrMetAlaPheLeuCys 28
DB 7 CGCAAAAAATTCGCTATTGGTCTGGCATGTCGGCGCACGCTCGGCATCTCGCC 66
QY 29 SerLeuArgGluLeuGlyValValLeuPheAspValProMetGly 48
DB 67 AGCTTGAGGAACCTGGCGATATCTCTTCGACATCGCGACGGATCCCGCAGGC 126
QY 49 LysAlaMetAspIleSerHisSerValValAspThrGlyIleThrValTyrGly 68
DB 127 AAGGTCTGGATATTGCCAGTCCGCGCGGTGAAGGCTTCAATGCAAAAGCTCTCCGC 186
QY 69 SerAsnSerTyrGluCysLeuLysGlyAlaAspValValIleThrAlaGlyIleThr 88
DB 187 GCTTCGGATTACCGCGCCATCAAGGCGCAGACGCTCGCATGTCACCGAGTCTCGCC 246
QY 89 LysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProValAsnIle 108
DB 247 CCAGACCCCGC-----ATGACCGCGGATGATCTTTCGCGCATCAACCTC 291
QY 109 LysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAlaPheValIle 128
DB 292 AAGGTTCATGAACAGTCCGCGCGCATCAAGAAATATGCTCCGAACGCTTTCGTGATC 351
QY 129 AsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeuPro 148
DB 352 TGCATCACCACACCGCTCGACGCCATGGTCTGGCGGCTGCAAGATCTTCGCGCCTGCGG 411
QY 149 HisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMetIle 168
DB 412 AAGAACAGGTCTGGCATGCTGGCGTTCTCGACGCGCAGCTTCGCCCTGTTCCTT 471
QY 169 AlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHisGly 188
DB 472 CGCGAAGATTCACAGTTTCGGTCCAGGACGTCACCGCTTCGTCGCGGTTCATGCG 531
QY 189 AspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGluPhe 208
DB 532 GACACCATGTCGCGCTCGCGTTATTCACCGTTGGCGGCTTCGCTTACCGATCTC 591
QY 209 ValLysLysGlyTrpIleLysGlnGluValAspAspIleValGlnLysThrLysVal 228
DB 592 GTCAAGATCGGTGGTTGACCGCGAACGCTTGACAGCATATCCAGCGCACCGCTGAC 651
QY 229 AlaGlyGlyLeuValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaProGlyAla 248
DB 652 GCGCGCGCGAAATCGTCGCGCTTTGAAGACCGGCTCGGCTTATTCGCGCGCGCT 711
QY 249 SerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCysSer 268

DB 712 TCGGCTATCGAATGGCGCAATCTACTCTCAAGACAAGAGCGGTTCGCCCGCTGCT 771
QY 269 CysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysValIle 288
DB 772 GCCACACCTTTCGGGCCAGTATGGCGTAGACACATATGATGTGCGCGTCCCGCATCATC 831
QY 289 GlyGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGluArgGlnGlu 308
DB 832 GGTCCGCGCGGTATCGAGCGCGTCTAGATCGAATCTGAACAAGAGAAAGAAGCCGCC 891
QY 309 LeuGlnGlySerIleAspGluValLysGluMetGlnLysAla 322
DB 892 TTCCAGAAATCCGTCGGCGCTGCTCGTGGTCTTTCCGAAGCC 933

RESULT 8

US-10-369-493-38561
; Sequence 38561, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 38561
; LENGTH: 960
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-38561

Alignment Scores:
Pred. No.: 1,93e-92 Length: 960
Score: 873.50 Matches: 167
Percent Similarity: 72.3% Conservative: 60
Best Local Similarity: 53.2% Mismatches: 82
Query Match: 51.7% Indels: 5
DB: 6 Gaps: 1

US-09-390-846-2 (1-330) x US-10-369-493-38561 (1-960)

QY 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyGlyThrMetAlaPheLeuCys 28
DB 7 CGCAAAAAATTCGCTATTGGTCTGGCATGTCGGCGCACGCTCGGCATCTCGCC 66
QY 29 SerLeuArgGluLeuGlyValValLeuPheAspValProMetGly 48
DB 67 AGCTTGAGGAACCTGGCGATATCTCTTCGACATCGCGACGGATCCCGCAGGC 126
QY 49 LysAlaMetAspIleSerHisSerValValAspThrGlyIleThrValTyrGly 68
DB 127 AAGGTCTGGATATTGCCAGTCCGCGCGGTGAAGGCTTCAATGCAAAAGCTCTCCGC 186
QY 69 SerAsnSerTyrGluCysLeuLysGlyAlaAspValValIleThrAlaGlyIleThr 88
DB 187 GCTTCGGATTACCGCGCCATCAAGGCGCAGACGCTCGCATGTCACCGAGTCTCGCC 246
QY 89 LysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProValAsnIle 108
DB 247 CCAGACCCCGC-----ATGACCGCGGATGATCTTTCGCGCATCAACCTC 291
QY 109 LysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAlaPheValIle 128
DB 292 AAGGTTCATGAACAGTCCGCGCGCATCAAGAAATATGCTCCGAACGCTTTCGTGATC 351

QY 129 AenileThrAsnProLeuAspValMetValAlaLeuGlnGluSerSerGlyLeuPro 148
 DB 332 TCATCACCAACCGCTGACCCCATGCTCGGGCGCTGCAAGATTCTCGGGCTGCGG 411
 QY 149 HisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMetIle 168
 DB 412 AAGAACAGGTGCTGGCATGCTGGCTGCTCGACAGGACGCTTCGGCTGTCTCT 471
 QY 169 AlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHisGly 188
 DB 472 GCCGAAGAATTCAACGTTTCGGTCCAGGACGCTCACCGCTTCTCTCGGGCTCATGGC 531
 QY 189 AspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGluPhe 208
 DB 532 GACACATGGTGGCGCTGCGCGTTATTCACCGCTTGGCGGCTTCCGCTTACCGATCTC 591
 QY 209 ValLysLysGlyTyrIleLysGlnGluValAspAspIleValGlnLysThrLysVal 228
 DB 592 GTCAAGATGGGTGCTGACCGCAACGCTTGTGAGACGATCATCCAGGACCGCTGAC 651
 QY 229 AlaGlyGlyIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaProGlyAla 248
 DB 652 GCGCGCGCGGAATCGTGGCGCTTGAAGACCGGCTCGGCTATTACGCGCGCGCTGCT 711
 QY 249 SerAlaIleGlnMetAlaGluSerTyrLysLysAspArgLysArgValMetValCysSer 268
 DB 712 TCGGCTATCGAATGCGCGAATCTTACCTCAAGGACAAAGACGCGGCTTCTGCGCGCTGCT 771
 QY 269 CysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysValIle 288
 DB 772 GCCACACCTTCGGGCGAGTATGGCTAGACGACATGTATGTCGGGCTGCCACCATCATC 831
 QY 289 GlyGlyArgGlyValGluLysIleLeuLeuGluLeuThrAlaGlnGluArgGlnGlu 308
 DB 832 GGTGCGCGCGTATCGAGCGGCTCATCGAGTCGACTGCAACAGGACAAAGACCGCC 891
 QY 309 LeuGlnGlySerIleAspGluValLysGluMetGlnLysAla 322
 DB 892 TTCAGAAATCGTGGCGCTGCTCGCTGCTTTCGGAAGCC 933

RESULT 9

US-10-369-493-44435
 ; Sequence 44435, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; PRIORITY FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 44435
 ; LENGTH: 933
 ; TYPE: DNA
 ; ORGANISM: Rhodopsedomonas palustris
 US-10-369-493-44435

Alignment Scores:
 Pred. No.: 3,67e-91 Length: 933
 Score: 862.50 Matches: 169
 Percent Similarity: 72.6% Conservative: 59
 Best Local Similarity: 53.8% Mismatches: 81
 Query Match: 51.1% Indels: 5
 DB: 6 Gaps: 1

US-09-390-846-2 (1-330) x US-10-369-493-44435 (1-933)

QY 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
 DB 7 CGCATATAAATTGCTTGGTCCGCCAGATCGCGGTACGCTGCACATTGGTC 66
 QY 29 SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMetGly 48
 DB 67 GGGTGAAGAGCTCGCGACGCTGTGTGTTCACATTGCGAGGGGCTGCGCAGAGGC 126
 QY 49 LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyrGly 68
 DB 127 AAGCGCTCGACATCGCGGAGTCTCGCGGTCGACGTTTCGACTCCAAGTCGACCGT 186
 QY 69 SerAsnSerTyrGluCysLeuLysGlyAlaAspValValIleIleThrAlaGlyIleThr 88
 DB 187 GCCAACTCTACGAGGCGATCGAAGGCGCGCTGGTGTGATCGTCACCGCGCGTGGCG 246
 QY 89 LysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProValAsnIle 108
 DB 247 CGCAAGCGCGGC-----ATGAGCGGTGACGATCTGCTCAGCATCAACCTC 291
 QY 109 LysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAlaPheValIle 128
 DB 292 AAGTGTATGGAGCAGTTCGGCGCTGGCATCAAGAGTACGCGCCGCGCGCTTCTGTATC 351
 QY 129 AsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeuPro 148
 DB 352 TGCATCACCAACCGCTCGACGCGATGCTGCGGCGCTGCAGAGGCGCTCGGCGCTGCGG 411
 QY 149 HisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMetIle 168
 DB 412 GCCAAGAGGTGTGCGCATGCGCGGCTGCTCGACTCGCGCGGCTTCCGTTACTTCCTG 471
 QY 169 AlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHisGly 188
 DB 472 GCCAGCAGTTCAACGTTGCTCGTCAAGACGCTACCGCTTTCGCTCGCGCGCCACCGT 531
 QY 189 AspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGluPhe 208
 DB 532 GATACCATGTTGCGCGCTGGTGAAGTACTCGACCGCTCGCGGAATTCCGCTCGCGCACCTG 591
 QY 209 ValLysLysGlyTyrIleLysGlnGluValAspIleValGlnLysThrLysVal 228
 DB 592 GTCAAGATGGGCTCGACCTCGCAGCGCGCTCGATGAGATCGTCGACCGCGCACCGCAAC 651
 QY 229 AlaGlyGlyIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaProGlyAla 248
 DB 652 GCGCGCGCGAGATCGTCAACCTGCTGAAGACCGGCTCGGCGTTCCTACGCGCGCGCTG 711
 QY 249 SerAlaIleGlnMetAlaGluSerTyrLysLysAspArgLysArgValMetValCysSer 268
 DB 712 TCGGCGATCGCATGCGCGAGAGTATCTGAGGACAAAGACGCGGTGGTCCCGTCCG 771
 QY 269 CysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysValIle 288
 DB 772 GCCACCTCAAGCGCAATACGCGCTCAAGACATGTATGTCGGGCTTCCGCTGCTGATC 831
 QY 289 GlyGlyArgGlyValGluLysIleLeuLeuGluLeuThrAlaGlnGluArgGlnGlu 308
 DB 832 GCGCACAAAGGCGTCGAGCGCATCGTCGAGATCGAGCTGCGCGCAAGGACAAAGGACG 891
 QY 309 LeuGlnGlySerIleAspGluValLysGluMetGlnLysAla 322
 DB 892 TTCAGACAGTTCGCTCGCTGCTTTCAGGGCTGCTGCGAGGCC 933

RESULT 10

US-10-369-493-35747
 ; Sequence 35747, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.


```
Db      892 TTCACAAAGTCGGTCAATCGGTCGAGGGCCTCGTCGAGGGCTCGAAGGGCATCGACTCC 951
Qy      329 Ser 329
      |||
Db      952 TCG 954

RESULT 13
US-10-369-493-34210
; Sequence 34210, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 34210
; LENGTH: 960
; TYPE: DNA
; ORGANISM: Sphingomonas aromaticivorans
US-10-369-493-34210

Alignment Scores:
Pred. No.: 2,24e-89 Length: 960
Score: 847.50 Matches: 167
Percent Similarity: 71.0% Conservative: 61
Best Local Similarity: 52.0% Mismatches: 88
Query Match: 50.2% Indels: 5
Gaps: 1

US-09-390-846-2 (1-330) x US-10-369-493-34210 (1-960)

Qy      9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
Db      7 CGCAAGAGATCGCGCTGATCGCGCGGCAACATCGGTGGCATCTTGGCCACCTCGCC 66

Qy      29 SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMetGly 48
Db      67 GCCCAGAAGAACTTTGGCGACATCTCTCTGTCGACGTCGTTGAAGGCGTGCCTCCAGGCG 126

Qy      49 LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyrgly 68
Db      127 AAGCGCGCTCGACCTTTTCGACGTGCGGTCCGGTCGGAAGGCTTCGACGCCCAACATCATCGCC 186

Qy      69 SerAsnSerTyrglyCysLeuLysGlyAlaAspValValIleIleThrAlaGlyIleThr 88
Db      187 ACCACAGCTACAAAGGATCCGACGCGGACGTCATCATCTGTCACCGCGCGTCCGCC 246

Qy      89 LysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProValAsnIle 108
Db      247 CGCAAGCCCGGC-----ATGAGCCGCGACGACCTGCTCGGCATCAACCTC 291

Qy      109 LysIleMetArgGluValGlyAlaIleLysSerTyrglyProAsnAlaPheValIle 128
Db      292 AAGGTGATGAAGCGCGTCCGCGAAGGATCCGCGCAACACGCGCGCATCGTTCGTGATC 351

Qy      129 AsnIleThrAsnProLeuAspValMetValAlaLeuGlnGluSerSerGlyLeuPro 148
Db      352 TCATCATCAACCAACCCGCTGACCGCATGCTCGGGCGCTGCGGAGTCTCGGCGCTCGCC 411

Qy      149 HisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMetIle 168
Db      412 GCGAACAAGGTCGTCGGCATGCCGCGGTCGACTCGGCGCGTCTCAGCACGTTCTCTG 471

Qy      169 AlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHisGly 188
```

```
Db      472 GCATGGGAATTCGGCGTCTCGATCCGCGACGTCGAACACAGCTTCGTCTCGCGGCCACCGC 531
Qy      189 AspHisMetValProLeuSerArgTyrglyAlaThrValAsnGlyIleProLeuSerGluPhe 208
Db      532 GACACCATGGTTCGGGTCAACCCAGTACTCGACCGTCAACCGCATCCCGGTGCGCGCATC 591

Qy      209 ValLysLysGlyTrpIleLysGlnGluValAspAspIleValGlnLysThrLysVal 228
Db      592 GTCAAGATGGCGCTGTCACCCAGGAAAGATCGACGCGATCGTCGACGCGCACCGCTCG 651

Qy      229 AlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrglyAlaProGlyAla 248
Db      652 GCGCGCGCGGATCGTCGGGCTGCTCAAGACCGGTTCCGGCTTCTACGCGCTCGCGCA 711

Qy      249 SerAlaIleGlnMetAlaGluSerTyrglyLeuLysAspArgLysArgValMetValCysSer 268
Db      712 TCGGCGATCGCATGGCGAAGCTATCTCAACGACCCAGAACGCGATCTCGCCCTCGGCC 771

Qy      269 CysTyrglyGlnGlyGlnTyrglyValGlnAsnHisTyrglyLeuGlyValProCysValIle 288
Db      772 GCCTACGTTGACGCGCAATACGCGGTGAACGGTCTTTACGTCGGCGTCCGGTCTGATC 831

Qy      289 GlyGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGluArgGlnGlu 308
Db      832 GCGCGGAACGCGTCGAGAAGGTGATCGAGATCGAACTCGACGACGAAGCAAGGCAAC 891

Qy      309 LeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaLeuAspAla 328
Db      892 CTCAGGTCTCCGTCGACGCGGTCAAGGAACTGCTGGAAGCCTGCAAGGGCATCGACCCC 951

Qy      329 Ser 329
      |||
Db      952 AGC 954

RESULT 14
US-10-369-493-46924
; Sequence 46924, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 46924
; LENGTH: 939
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-10-369-493-46924

Alignment Scores:
Pred. No.: 1.03e-76 Length: 939
Score: 740.00 Matches: 147
Percent Similarity: 65.8% Conservative: 63
Best Local Similarity: 45.9% Mismatches: 98
Query Match: 43.8% Indels: 12
Gaps: 3

US-09-390-846-2 (1-330) x US-10-369-493-46924 (1-939)

Qy      7 AsnThrArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPhe 26
Db      7 AATACTCGTAAAAAGTTTCTGTTATCGGAGCAGGTTTACCGGAGCTTACCACTCATTT 66
```

```

QY 27 LeuCysSerLeuArgGluLeuGlyAspValValLeuPheAspVal-----ValProAsn 44
DB 67 TTAATCGCTCAAAAGAGCTGGCAGACGTTGTTCTTGTGACATTCGCGAAATTTGGAGAAC 126
QY 45 MetProMetGlyLysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIle 64
DB 127 CGGACAAAGGAAAGCGCTTGATGCTTGAAGCAAGCCGGTTCAAGGCTTTGACGCA 186
QY 65 ThrValTyGlySerAsnSerTyGluCysLeuLeuGlyAlaAspValValIleThr 84
DB 187 AAATATTACGGGAACATCCCAATTACGAGGATACAGCCGGCTCTGACATTTGTCATTACA 246
QY 85 AlaGlyIleThrLysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeu 104
DB 247 GCGGGTATCGCAAAACCTGGT-----ATGAGCAGAGATGATCTGGTC 291
QY 105 ProValAsnIleLysIleMetArgGluValGlyAlaAlaIleLysSerTyCysProAsn 124
DB 292 TCTACAAACGAAAGATTATGAGAAGCGTTACCGAGGAATCTGGAATATTCTCTGAC 351
QY 125 AlaPheValIleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSer 144
DB 352 TCTATTATTGTGCTGCAAAATCTGTTGATGCAATGACATACCGCGTGTACAAGAA 411
QY 145 SerGlyLeuProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPhe 164
DB 412 TCAGGCTTCCCTAAAGAGCGGTGTAATCGGCCAGTACAGGTGCTGATACGCGGATTC 471
QY 165 ArgArgMetIleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIle 184
DB 472 AGNACATTTGTGGCAGAGAAATTAACCTGTCTGTAAGATGTGATGTTGTTGTCATC 531
QY 185 GlyValHisGlyAspHisMetValProLeuSerArgTyAlaThrValAsnGlyIlePro 204
DB 532 GCGCGACACGGTCACGATATGTTGCGCTTGTGCGTTATCTTATGCTGCGGTATCCCG 591
QY 205 LeuSerGluPheValLysLysGlyTrpIleLysGlnGluValAspAspIleValGln 224
DB 592 CTTGAAACTCTTATCCGAAA-----GAAAGGATTGACCAATTTGTTGGAG 636
QY 225 LysThrLysValAlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyTrp 244
DB 637 CGCACTAGAAAGCGGCGGAAATCGTGAATCTTCTTGGAAACGGAAGCGCGTATTAT 696
QY 245 AlaProGlyAlaSerAlaIleGlnMetAlaGluSerTyLeuLysAspArgLysArgVal 264
DB 697 GCGCTGCGGCTTCTCTGACAGAAATGTCGAAGCGATCTTGAAGATCAGCGCGGTC 756
QY 265 MetValCysSerCysTyTrpLeuGlnGlyGlnTyGlyValGlnAsnHisTyTrpLeuGlyVal 284
DB 757 CTTCTACAAATTGCTTATCTTGAAGGGAATACGGCTATCAAGGCACTCTACCTTGGTGT 816
QY 285 ProCysValIleGlyCysArgValGluLysIleIleGluLeuLeuThrAlaGln 304
DB 817 CTTACAAATTGTAGCGGCAACGCTCTTGAAGCAATCATTTGAATCTGAACCTGACAGACTAT 876
QY 305 GluArgGlnGlnGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAla 324
DB 877 GAAAGCGCAGCTGAATAAATCAGTTGATCTGTGCANAAATGTCATGAAGATTATATCC 936

```

RESULT 15

```

US-10-369-493-41082
; Sequence 41082, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B

```

```

; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 41082
; LENGTH: 945
; TYPE: DNA
; ORGANISM: Bacillus halodurans
US-10-369-493-41082

Alignment Scores:
Pred. No.: 4,68e-73 Length: 945
Score: 709.00 Matches: 140
Percent Similarity: 65.1% Conservative: 67
Best Local Similarity: 44.0% Mismatches: 97
Query Match: 42.0% Indels: 14
DB: 6 Gaps: 4

US-09-390-846-2 (1-330) x US-10-369-493-41082 (1-945)

```

```

QY 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyGlyThrMetAlaPheLeuCys 28
DB 13 CCGCAAAAGGTCCTCGTTATTGGGGCGGGTTTCACAGGTGCTACAGACGGCTAATGGTT 72
QY 29 SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMet----- 45
DB 73 GCTCAAAAGGAATTAGGTGATGCTGTTCTTGTGGAT---ATTCCGCAATGCAAGTCCG 129
QY 46 ProMetGlyLysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThr 65
DB 130 ACTAAAGGAAAGCGTTAGATATGTCGAGTCAACCGCGGTTCAAGGGGTTGATGTGAAC 189
QY 66 ValTyGlySerAsnSerTyGluCysLeuLysGlyAlaAspValValIleThrAla 85
DB 190 ATCAGAGTACGCTAGCTATGATATACGAGGATTCGATGTTGTTGTCATTACGGGT 249
QY 86 GlyIleThrLysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuPro 105
DB 250 GGAATTGCTCGTAAGCCTGGT-----ATGAGCGCGATGATCTCGTATCA 294
QY 106 ValAsnIleLysIleMetArgGluValGlyAlaAlaIleLysSerTyCysProAsnAla 125
DB 295 ACAATTCGGGAATTTATGAAGCTGTTACGAAAGAAAGTGGTGAACATTCACCAACGCC 354
QY 126 PheValIleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSer 145
DB 355 TACATTATGTTTAAACCAACCCACAGATGCGATGACCTATACCGTTTATAAAGATCT 414
QY 146 GlyLeuProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArg 165
DB 415 GGTTCCTCCGAAAAACCGGTCATCGTCAATCGGGGGTTCATAGATACGCGACGCTCCGT 474
QY 166 ArgMetIleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGly 185
DB 475 ACGTTCGTCGCGCAAGAGCTTAACCTATCTGTTGAAGATATTACAGGTTTGTCTTGGG 534
QY 186 ValHisGlyAspHisMetValProLeuSerArgTyAlaThrValAsnGlyIleProLeu 205
DB 535 GGGCAGCGGGATGACATGTCCTTGTGATTCGTTACTCTATGTCAGGTGGGATTCGGT 594
QY 206 SerGluPheValLysGlyTrpIleLysGlnGluValAspAspIleValGlnLys 225
DB 595 GAAAGCTC-----CTTCCACAAAGAGCGCATCGATCGGATGTGGAACGT 639
QY 226 ThrLysValAlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyTrpAla 245
DB 640 ACGCGCAAAAGCGCGGTGAATTTGTCGGCTTGTCTGGAATGCGAGTCTTATTATGCA 699
QY 246 ProGlyAlaSerAlaIleGlnMetAlaGluSerTyLeuLysAspArgLysArgValMet 265
DB 700 CACGCGCTTCACTTGCAGAGATGTCGAAGCCATCTTGAAGATTAAGAAACGCGCTCT 759

```

```

Qy 266 ValCysSerCysTyrLeuGlnGlyGlnTyrGlyValGlnAenHisTyrLeuGlyValPro 285
Db 760 CCAACGATTGCTTACCTTGAAGGTGAGTATGGATATGAAGATATCTATGTTGGAGTCCG 819
Qy 286 CysValIleGlyGlyArgGlyValGlyLysIleIleGluLeuGluLeuThrAlaGlnGlu 305
Db 820 ACGATTCTTGGCGGAGACGCCATTGANAAGGTTATTGAGCTTGATCTAACCCGATGAAGAG 879
Qy 306 ArgGlnGluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIle 323
Db 880 AAAGCGACCTTGTGCTAAGTCCATTGAATCTGTTCGCAACGTCATGTCGGCACTA 933

```

Search completed: March 3, 2006, 00:57:35
Job time : 950 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 3, 2006, 00:57:49 ; Search time 405 Seconds
(without alignments)
1786.381 Million cell updates/sec

Title: US-09-390-846-2

Perfect score: 1688

Sequence: 1 MAVPEKTRPKIAVGVSGMI.....GSIDVEKEMQKAIALDASK 330

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7218512 seqs, 1096188070 residues

Total number of hits satisfying chosen parameters: 14437024

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abss/ABSSWEB.spool/US09390846/runat 02032006 104245 5144/app query.fasta 1
-DB=Published Applications NA New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs03h
-USER=US09390846 @CGN 1 1 220 @runat 02032006 104245 5144 -NCPU=6 -ICPU=3
-NO_MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELEXT=7

Database :

Published Applications NA New:*

1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	880.5	52.2	2862	12 US-11-156-953-6	Sequence 6, Appli
2	880.5	52.2	4716	12 US-11-156-953-4	Sequence 4, Appli
3	460	27.3	1363	9 US-11-077-619-51	Sequence 51, Appli
4	418.5	24.8	969	12 US-11-074-176-193	Sequence 193, App

5	408	24.2	1661	8 US-10-947-249-109	Sequence 109, App
6	408	24.2	1661	12 US-11-000-688-1061	Sequence 1061, Ap
7	406	24.1	1336	8 US-10-955-054A-92	Sequence 92, Appl
8	406	24.1	1336	7 US-10-947-249-93	Sequence 93, Appl
9	405	24.0	1171	7 US-10-506-443A-1	Sequence 1, Appli
10	401.5	23.8	969	8 US-10-793-626-1411	Sequence 1411, Ap
11	401.5	23.8	969	8 US-10-793-626-2383	Sequence 2383, Ap
12	401.5	23.8	3019	8 US-10-793-626-4355	Sequence 4355, Ap
13	401.5	23.8	4093	8 US-10-793-626-3793	Sequence 3793, Ap
14	396	23.5	1317	12 US-11-128-061-676	Sequence 676, App
15	396	23.5	1317	12 US-11-128-049-676	Sequence 676, App
16	356	21.1	1656	5 US-09-978-360A-265	Sequence 265, App
17	338.5	20.1	1053	7 US-10-506-443A-2	Sequence 2, Appli
18	338	20.0	1141	6 US-09-925-065A-722171	Sequence 722171,
19	335.5	19.9	1047	7 US-10-506-443A-5	Sequence 5, Appli
20	285.5	16.9	1280	12 US-11-136-527-1594	Sequence 1594, Ap
21	250.5	14.8	879	7 US-10-506-443A-3	Sequence 3, Appli
22	240.5	14.2	953	12 US-11-128-061-348	Sequence 348, App
23	240.5	14.2	953	12 US-11-128-049-348	Sequence 348, App
24	228	13.5	811	7 US-10-506-443A-4	Sequence 4, Appli
25	192.5	11.4	600	12 US-11-128-061-3990	Sequence 3990, Ap
26	192.5	11.4	600	12 US-11-128-049-3990	Sequence 3990, Ap
27	182	10.8	391	8 US-10-821-234-394	Sequence 394, App
28	162.5	9.6	1225	6 US-09-925-065A-26579	Sequence 26579, A
29	161.5	9.6	1268	12 US-11-000-688-1283	Sequence 1283, Ap
30	140	8.3	1307	12 US-11-136-527-38	Sequence 38, Appl
31	129.5	7.7	600	12 US-11-128-061-4318	Sequence 4318, Ap
32	129.5	7.7	600	12 US-11-128-049-4318	Sequence 4318, Ap
33	128.5	7.6	1637	12 US-11-136-527-1710	Sequence 1710, Ap
34	122.5	7.3	462	12 US-11-108-172-164	Sequence 164, App
35	110	6.5	1400	12 US-11-136-527-5806	Sequence 5806, Ap
36	108.5	6.4	412	12 US-11-108-172-652	Sequence 652, App
37	107	6.3	492	5 US-09-995-493-195	Sequence 195, App
38	106	6.3	389	8 US-10-821-234-381	Sequence 381, App
39	94.5	5.6	861	8 US-10-467-657-2767	Sequence 2767, Ap
40	93	5.5	1336	12 US-11-195-459-24	Sequence 24, Appl
41	92.5	5.5	1362	12 US-11-074-176-141	Sequence 141, App
42	92	5.5	1605	8 US-10-646-283-1	Sequence 1, Appli
43	91	5.4	1065	8 US-10-793-626-3097	Sequence 3097, Ap
44	91	5.4	3320	8 US-10-793-626-3649	Sequence 3649, Ap
45	91	5.4	3690	8 US-10-793-626-3439	Sequence 3439, Ap

ALIGNMENTS

RESULT 1
US-11-156-953-6
; Sequence 6, Application US/11156953
; Publication No. US20060003051A1
; GENERAL INFORMATION:
; APPLICANT: Cleary, Joseph
; APPLICANT: Coleman, Russell
; APPLICANT: Harding, Nancy
; APPLICANT: Patel, Yamini
; TITLE OF INVENTION: GENETICALLY PURIFIED GELLAN GUM
; FILE REFERENCE: 012047.00029
; CURRENT APPLICATION NUMBER: US/11/156.953
; CURRENT FILING DATE: 2005-06-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 2862
; TYPE: DNA
; ORGANISM: Spingomonas elodea
US-11-156-953-6

Alignment Scores:			
Pred. No.:	2,86e-94	Length:	2862
Score:	880.50	Matches:	173
Percent Similarity:	70.7%	Conservative:	54
Best Local Similarity:	53.9%	Mismatches:	89
Query Match:	52.2%	Indels:	5
DB:	12	Gaps:	1

```
US-09-390-846-2 (1-330) x US-11-156-953-6 (1-2862)
; Sequence 4, Application US/11156953
; Publication No. US2006003051A1
; GENERAL INFORMATION:
; APPLICANT: Cleary, Joseph
; APPLICANT: Coleman, Russell
; APPLICANT: Harding, Nancy
; APPLICANT: Patel, Yamin
; TITLE OF INVENTION: GENETICALLY PURIFIED GELLAN GUM
; FILE REFERENCE: 012047.00029
; CURRENT APPLICATION NUMBER: US/11/156,953
; CURRENT FILING DATE: 2005-06-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 4716
; TYPE: DNA
; ORGANISM: Spingomonas elodea
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1591)...(3450)
; US-11-156-953-4

Alignment Scores:
Pred. No.: 5.39e-94 Length: 4716
Score: 880.50 Matches: 173
Percent Similarity: 70.7% Conservative: 54
Best Local Similarity: 53.9% Mismatches: 89
Query Match: 52.2% Indels: 5
DB: 12 Gaps: 1

US-09-390-846-2 (1-330) x US-11-156-953-4 (1-4716)
Qy 9 ArgProlylAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
Db 3735 CGCAAGAGATCGCGCTCGATCGGTCGCGATGATCGGTCGCGACCTTCGCT 3794
Qy 29 SerLeuArgGluLeuGlyAspValValPheAspValProAsnMetProMetGly 48
Db 3795 GCATCAAGNACTGGCGGATATCGTCTGTTGCACATTGCGAGGGCAGCCGCGAGGC 3854
Qy 49 LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyGly 68
Db 3855 AAGCGCTCGACATCGCACAGTTCGCGACCGGTTCGAGGCTTCGATGCGCAACCTCAAGGCG 3914
Qy 69 SerAsnSerTyGluCysLeuLysGlyAlaAspValValIleIleThrAlaGlyIleThr 88
Db 3915 GCGAACAGCTACAGGACATTCGCGGCGCGACGTCGTCATCGTCACCGCGGTATCCG 3974
Qy 89 LysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProValAsnIle 108
Db 3975 CGCAAGCGGCGC-----ATGAGCGCGCGACGATCTTCTCAAGACCAACCTG 4019
Qy 109 LysIleMetArgGluValGlyAlaAlaIleLysSerTyCysProAsnAlaPheValIle 128
Db 4020 GGCGTGATGAAGCGCTGGCGAGGCGATCGCGCGACGCGCGCGCGCGCGCGTTCGTGATC 4079
Qy 129 AsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeuPro 148
Db 4080 TGCATACCAACCGCGCTCGACGCGATGTTGGGCGCTGCGCGGCTTCCTGGGCGCTGCG 4139
Qy 149 HisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMetIle 168
Db 4140 CACCAAGAGTCTCGCGCATGCGCGCGCTGTCGACTCGGCGCGCTTCGCCACCTTCCTC 4199
Qy 169 AlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHisGly 188
Db 4200 GCCGAGGAATCAACGCTCTCGGTGAGGACGTCACCGCTTCGTGCTGCGCGGACACGCG 4259
Qy 189 AspHisMetValProLeuSerArgTyAlaThrValAsnGlyIleProLeuSerGluPhe 208
Db 4260 GACACCATGGTCCCGGTGATCGAGTATTCGACCGCTCGCGCGGATCCCGGTGCGCGACCTG 4319
Qy 209 ValLysLysGlyTrpIleLysGlnGluValAspIleValGlnLysThrLysVal 228

US-09-390-846-2 (1-2862)
Qy 9 ArgProlylAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
Db 1881 CGCAAGAGATCGCGCTCGATCGGTCGCGATGATCGGTCGCGACCTTCGCTCGCT 1940
Qy 29 SerLeuArgGluLeuGlyAspValValPheAspValValProAsnMetProMetGly 48
Db 1941 GCAATCAAGAACTGGCGGATATCGTCTGTTGCACATTGCGAGGGCAGCCCGAGGCG 2000
Qy 49 LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyGly 68
Db 2001 AAGCGCTCGACATCGCACAGTTCGCGCGCGTTCGAGGCTTCGATGCGCAACCTCAAGGC 2060
Qy 69 SerAsnSerTyGluCysLeuLysGlyAlaAspValValIleIleThrAlaGlyIleThr 88
Db 2061 CGCAACAGCTACGAGGACATTCGCGCGCGCGCGCGCTGCTGATCGTCACGCGCGTATCCG 2120
Qy 89 LysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProValAsnIle 108
Db 2121 CGCAAGCGGCGC-----ATGAGCGCGCGACGATCTTCTCAAGACCAACCTG 2165
Qy 109 LysIleMetArgGluValGlyAlaAlaIleLysSerTyCysProAsnAlaPheValIle 128
Db 2166 GCGGTGATGAAGCGCTGGCGAGGCGATCGCGCGCACGCGCGCGCTTCGTGATC 2225
Qy 129 AsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeuPro 148
Db 2226 TGCATCAACCAACCGCTCGACGCGATGTTGGGCGCTGCGCGATTCGCGCGCTGCGC 2285
Qy 149 HisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMetIle 168
Db 2286 CACCAAGAGTCTCGCGATGCGCGCGCTGCTCGACTCGCGCGCTTCGCGACCTTCCTC 2345
Qy 169 AlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHisGly 188
Db 2346 GCCGAGGAATTCACGCTTCGTCGAGGACGTCACCGCTTCGTCGCGCGCGCGCGCGCG 2405
Qy 189 AspHisMetValProLeuSerArgTyAlaThrValAsnGlyIleProLeuSerGluPhe 208
Db 2406 GACACCATGGTCCCGGTGATCGAGTATTCGACCGTTCGCGCGCGCTTCGCGCGCGCG 2465
Qy 209 ValLysLysGlyTrpIleLysGlnGluValAspIleValGlnLysThrLysVal 228
Db 2466 ATCAAGATGGGTGCTCCACCGAGCGCATCGACCGCATCGTCGCGCGCGCGCGCGCTG 2525
Qy 229 AlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyTrpAlaProGlyAla 248
Db 2526 GCGCGCGCGGAGATCGTCGCTGCTCAAGACCGCTCGGCTATATATGCGCGCGCGCG 2585
Qy 249 SerAlaIleGlnMetAlaGluSerTyLeuLysAspArgLysArgValMetValCysSer 268
Db 2586 AGCGCGATCGCGATGCGCGAGAGTATCTGAAGGACAAAGAGCGCTTCCTTCCTGTGCC 2645
Qy 269 CysTyTrpLeuGlnGlyGlnTyTrpGlyValGlnAsnHisTyLeuLysValProCysVal 288
Db 2646 GCGCACCTCACCGCGCGTACGCGCTCGACGATCTGTACGTCGCGCGCTGCGCGCTGATC 2705
Qy 289 GlyGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGluArgGlnGlu 308
Db 2706 GCGAAGGACGCGCTCGAGCGCATCTCGAGATCGAGCTGAACCGCGCGCGAGAGAGAT 2765
Qy 309 LeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaLeuAspAla 328
Db 2766 TTCGAGCTCTCGTTCGATCGGTCAAGGAACCTGGTCGCGAGCATCGAAGTTCGATCGAC 2825
Qy 329 Ser 329
Db 2826 TCG 2828

RESULT 2
US-11-156-953-4
```

```

Db      4320 ATCAAGATGGGTGGTCCACCAGGAGCGCATCGACCCATCGTCCGGCGCACCGCTCG 4379
Qy      229 AlaGlyGlyGluValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaProGlyAla 248
Db      4380 GCGCGCGCGGAGATCGTCGCGTCTCAAGACCGGCTCGGCTATATTATGCGCGCGCAC 4439
Qy      249 SerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCysSer 268
Db      4440 ACGCGGATCGGATGGCGGAGAGCTATCTGAAGGACAAGACGCGCTCTCCCTGTGCG 4499
Qy      269 CysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysValIle 288
Db      4500 GCGCACTCACCGGCAGTACCGGCTCGAGCATCTGTAGCTCGGCTGCGGATCGTCAATC 4559
Qy      289 GlyGlyArgGlyValGluLysIleLeuLeuGluLeuThrAlaGlnGluArgGlnGlu 308
Db      4560 GCGAAGGACGGCGTCCGAGCGCATCTCGAGATCGAGCTGAACGCCACCGCGAAGCAGAT 4619
Qy      309 LeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaLeuAspAla 328
Db      4620 TTCGACGCTCGGTCTGATCGGTCAAGGAACTGGTCCGACGATGCAAGTCTGATCGACGCC 4679
Qy      329 Ser 329
Db      4680 TCG 4682

```

RESULT 3

```

US-11-077-619-51
; Sequence 51, Application US/11077619
; Publication No. US2006040279A1
; GENERAL INFORMATION:
; APPLICANT: Feesche, Joerg
; APPLICANT: Maurer, Karl-Heinz
; APPLICANT: Breves, Roland
; APPLICANT: Schweder, Thomas
; APPLICANT: Hecker, Michael
; APPLICANT: Juergen, Britta
; APPLICANT: Voigt, Birgit
; TITLE OF INVENTION: DNA CHIPS USED FOR BIOPROCESS CONTROL
; FILE REFERENCE: HENK-0122 / H5692
; CURRENT APPLICATION NUMBER: US/11/077,619
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/009979
; PRIOR FILING DATE: 2003-09-09
; PRIOR APPLICATION NUMBER: DE 10242433.0
; PRIOR FILING DATE: 2002-09-11
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 51
; LENGTH: 1363
; TYPE: DNA
; ORGANISM: Bacillus subtilis
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(1363)
; OTHER INFORMATION: ldn
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1166)
US-11-077-619-51

```

Alignment Scores:

```

Pred. No.: 1,78e-44 Length: 1363
Score: 460.00 Matches: 106
Percent Similarity: 55.2% Conservative: 69
Best Local Similarity: 33.4% Mismatches: 124
Query Match: 27.3% Indels: 18
DB: Gaps: 7

```

US-09-390-846-2 (1-330) x US-11-077-619-51 (1-1363)

Qy 11 LysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCysSerLeu 30

```

Db      222 AAAGTAGCTTTAATCGGAGCGGTTTGTGTGAACAGTATTGCGTATTGCGTAAATTAAC 281
Qy      31 ArgGluLeuGlyAsp---ValValLeuPheAspValValProAsnMetProMetGlyLys 49
Db      282 CAAGGATACACAGATGAGCTTGTGCTCATGTATGATGATAAATAAGAAAAAGCAATGGCGAT 341
Qy      50 AlaMetAspIleSerHisAsnSerSerValValAspThrGlyIle-----Thr 65
Db      342 GTGATGATTTTACCCACGGAAGCGTTT-----GGCTACAACCGGTCAAAAACA 392
Qy      66 ValTyrGlySerAsnSerTyrCysLeuLysGlyAlaAspValValIleIleThrAla 85
Db      393 TCTTACGGA-----ACATATGAAGACTGCAAGGATGCTGATATTGTCTGCAATTCGCGC 446
Qy      86 GlyIleThrLysIleProGlyLysSerAspLysGluTyrSerArgMetAspLeuLeuPro 105
Db      447 GGAGCAACCAACCAAACTGGTGAG-----ACACGCTTGAATTAGTAGNA 491
Qy      106 ValAsnIleLysIleMetArgGluValGlyAlaIleLysSerTyrCysProAsnAla 125
Db      492 AAGAACTTGAAGATTTTCAAGGCACTGTTAGTGAAGTCATGGCGAGCGGATTTCACGCC 551
Qy      126 PheValIleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSer 145
Db      552 ATTTTCTTAGTCGCGACAAATCCGGTTGATATCTCTGACTTACGCAACATGGAAATTCAGC 611
Qy      146 GlyLeuProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArg 165
Db      612 GCGCTGCCAAAGAGCGGCTGATTGGAAGCGGCACACACTTGTATTCGCGAGATTCCT 671
Qy      166 ArgMetIleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGly 185
Db      672 TTCATGCTGAGCGAATACTTTGGCGCAGCGCTCAAAACGTACACCGCATATTATCGGA 731
Qy      186 ValHisGlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeu 205
Db      732 GAGCACGCGCACACAGACTTCTCTGTTGGAGCCACGCGAATGTCCGCGGTGTCCGCTC 791
Qy      206 SerGluPheValLysLysGly---TyrIleLysGlnGluValAspAspIleValGln 224
Db      792 AGTGAATCTGTTGAGAAAAACGATCGGTACAAACAGAGAGAGCTGGACCAATTTGTAGAT 851
Qy      225 LysThrLysValAlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrTyr 244
Db      852 GATGTGAAAAACGCGACTTACCATATCATTCAG-----AAAAAGGCGCGACTTATTAT 905
Qy      245 AlaProGlyAlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgVal 264
Db      906 GGGGTTCGATGAGTCTTGTCTCGCATTAACAAAAGCCATTCTTCATAATGAAAAACAGCATA 965
Qy      265 MetValCysSerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyVal 284
Db      966 TTAATCTGCACACATATTTGGACGGCAATACGCGTCAGATGCGTGTACATGCTGCTGTG 1025
Qy      285 ProCysValIleGlyGlyArgValGluLysIleIleGluLeuLeuLeuThrAlaGln 304
Db      1026 CCGGCTGTCGTGAATCGCGGAGGATCCAGGTATCCTGAGCTGAACTTAATGAGAAA 1085
Qy      305 GluArgGlnGluLeuGlnGlySerIleAspGluValLysGluMetGlnLys 321
Db      1086 GAAAAAGAACAGTTCCTTTCACAGCGCGCGCTCTTAAAAACATTTTAAAA 1136

```

RESULT 4

```

US-11-074-176-193
; Sequence 193, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Peril, Andrea Ascarate

```

```

; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 193
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Lactobacillus acidophilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(969)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: ORF 271; L-lactate dehydrogenase; L-LDH
US-11-074-176-193

```

```

Alignment Scores:
Pred. No.: 1,04e-39 Length: 969
Score: 418.50 Matches: 101
Percent Similarity: 56.3% Conservative: 83
Best Local Similarity: 30.9% Mismatches: 124
Query Match: 24.8% Indels: 19
DB: 12 Gaps: 10

```

```
US-09-390-846-2 (1-330) x US-11-074-176-193 (1-969)
```

```

Qy 1 MetAlaValPheGluLysAsnThrArgProLysIleAlaMetValGlySerGlyMetIle 20
Db 1 ATGGCAAGAGTTGAAACCTCGT-----AAAGTTATTTAGTTGGTGACGGTCTGTA 54
Qy 21 GlyGlyThrMetAlaPheLeuCysSerLeuArgGluLeuGly-----AspValVal 37
Db 55 GGTCTCTACCTTTGCAATTT-----TCAATGGTGCAACAGGTATTGCTGAAGAATTAGGT 108
Qy 38 LeuPheAspValProAsnMetProMetGlyLysAlaMetAspIleSerHisAsnSer 57
Db 109 ATCATTTGATTGCTTAAGAACACAGTTGNAGGTGACGATCGATTTAGCA---GATGCT 165
Qy 58 SerValValAspThrGlyIleThrValTyrglySerAsnSerTyrgluCysLeuLysGly 77
Db 166 ACTCCATGGACTTCCCAAGAACATTTTACGCAGCTGACTACGCTGACTGC---AAGGAC 222
Qy 78 AlaAspValValIleThrAlaGlyIleThrLysIleProGlyLysSerAspLysGlu 97
Db 223 GCAGACTTAGTAGTTATTACTGCTGCTGCCACAAAGCCAGGTGAA----- 270
Qy 98 TrpSerArgMetAspLeuLeuProValAsnIleLysIleMetArgGlu---ValGlyAla 116
Db 271 ---ACTGCTTGACCTGTTTAAACAGACTTGAAGATTTATCATCATCGTTGAACCA 327
Qy 117 AlaIleLysSerTyrglyProAsnAlaPheValIleAsnIleThrAsnProLeuAspVal 136
Db 328 GTTGTGTAATCAGGCTTTTGAAGGTATCTTCTTAGTA---GTTGCTAAACCCAGTTGACATC 384
Qy 137 MetValAlaAlaLeuGlnGlnSerSerGlyLeuProHisHisArgIleCysGlyMetAla 156
Db 385 TTGACTCACGCACTTTGGAAGATTTTCAAGCTTCCCTAAGGATCGCGTTATTGCTTCAGGT 444
Qy 157 GlyMetLeuAspSerSerArgPheArgMetIleAlaAspLysLeuGluValSerPro 176
Db 445 ACTTCACTTGATCTGCTGCTTCTTCAAAAGGTTATCGTTAAGTGAACACAGCTTGACCA 504
Qy 177 ArgAspValGlnGlyMetValIleGlyValHisGlyAspHisMetValProLeuSerArg 196
Db 505 CGTTCAAGTTAATGCATACATGCTTGTGTGAACACGGGTGACTGAATTTCCAGTATGGAGC 564
Qy 197 TyrAlaThrValAsnGlyIleProLeuSerGluPheValLysGlyTrpIleLysGln 216

```

```

Db 565 TACAACAATGTTGCTGGCGTAAAGGTTAGCGACTGGGTTAAGGCTCACGGTATGGATGAA 624
Qy 217 GluGluValAspAspIleValGlnLysThrLysValAlaGlyGlyGluIleValArgLeu 236
Db 625 TCTAAGCTTGAAGAAATCCACAAGGAAGTTGCTGCACATGGCTTACGACATTATCAAC--- 681
Qy 237 LeuGlyGlnGlySerAlaTyrglyAlaProGlyAlaSerAlaIleGlnMetAlaGluSer 256
Db 682 ---AAGAAGGGTGCTACTTTCTACGGTATCGGTATCGGTACAGCAATGATCGCTAAGGCT 738
Qy 257 TyrLeuLysAspArgLysArgValMetValCysSerCysTyrglyLeuGlnGlyGlnTyrgly 276
Db 739 ATCTTGAACGATGAACACCGGTACTTCCACTCTCAGTTGCAATGGATGGTCAATACGGT 798
Qy 277 ValGlnAsnHisTyrglyLeuGlyValProCysValleGlyGlyArgGlyValGluLysIle 296
Db 799 TTACACGACCTTCACATTTGGTACTCTCTCGCATGTTTGTGGCCGTAAACGGTCTTTGAACAAAT 858
Qy 297 IleGluLeuGluLeuThrAlaGlnGluArgGlnGluLeuGlnGlySerIleAspGluVal 316
Db 859 ATTGAATGCTTTTAAACCGCTGATGAACAAGCTAAGATGGAAGCTTCTGCTAAGCAATTA 918
Qy 317 LysGlu---MetGlnLysAla 322
Db 919 AAGGAGTTATGGACAAGGCC 939

```

RESULT 5

```

US-10-947-249-109
; Sequence 109, Application US/10947249
; Publication No. US20050287541A1
; GENERAL INFORMATION:
; APPLICANT: Akira NAKAGAWARA
; APPLICANT: Miki OHIRA
; APPLICANT: Shin ISHII
; APPLICANT: Takeshi GOTO
; APPLICANT: Hiroyuki KUBO
; APPLICANT: Takahiro HIRATA
; APPLICANT: Yasuko YOSHIDA
; APPLICANT: Saichi YAMADA
; TITLE OF INVENTION: Microarray for Predicting the Prognosis of Neuroblastoma
; FILE REFERENCE: 117007
; CURRENT APPLICATION NUMBER: US/10/947,249
; CURRENT FILING DATE: 2004-09-23
; PRIOR APPLICATION NUMBER: US 60/505,614
; PRIOR APPLICATION NUMBER: 2003-09-25
; NUMBER OF SEQ ID NOS: 200
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 109
; LENGTH: 1661
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-947-249-109

```

```

Alignment Scores:
Pred. No.: 3,71e-38 Length: 1661
Score: 408.00 Matches: 102
Percent Similarity: 53.6% Conservative: 71
Best Local Similarity: 31.6% Mismatches: 138
Query Match: 24.2% Indels: 12
DB: 8 Gaps: 7

```

```
US-09-390-846-2 (1-330) x US-10-947-249-109 (1-1661)
```

```

Qy 5 GluLysAsnThrArgProLysIleAlaMetValGlySerGlyMetIleGlyGlyThrMet 24
Db 143 GAACAGACCCCAAGATTAAGATTTGTTGGGTTGGTGTGTTGGCATGCGCTGT 202
Qy 25 AlaPheLeuCysSerLeuArgGluLeuGlyAsp---ValValLeuPheAspValValPro 43
Db 203 GCCATCATGTTCTTAATGAAGACTTGGCAGATGAACCTTGCTCTTTGTTGATGTCAGAA 262

```

QY 44 AsnMetProMetGlyLysAlaMetAspIleSerHisAsnSerSerValValAspThrGly 63
 DB 263 GACAAATGAAGGAGATGATGATCTCCAACTCCAGCGCTTTTCCTTAGAACACCA 322
 QY 64 IleThrValTyrGlySerAsnSerTyrGluCysLeuLysGlyAlaAspValValIleIle 83
 DB 323 AAGATTGCTCTGGCAAGAC---TATAATGTAACCTGCAAACTCCAGCTGTCATATC 379
 QY 84 ThrAlaGlyIleThrLysIleProGlyLysSerAspLysGluTyrSerArgMetAspLeu 103
 DB 380 ACGCTGGGGCAGCTCAGCAAGAGGAGAA-----AGCGCTCTTAATTG 424
 QY 104 LeuProValAsnIleLysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysPro 123
 DB 425 GTCCAGCGTAACTGAACATATTTAAATTCATCTCTTAATGTTGTAATAACAGCCG 484
 QY 124 AsnAlaPheValIleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGlu 143
 DB 485 AACTGCAAGTTGCTTATTGTTTCAATCCAGTGGATATCTTGACCTACGTGGCTTGAAG 544
 QY 144 SerSerGlyLeuProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerArg 163
 DB 545 ATAAGTGTGTTTCCCAAAACCGTGTATTGGAAGTGTGCAATCTGCAATCTGCAATCTG 604
 QY 164 PheArgArgMetIleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetVal 183
 DB 605 TTCCGTTACCTGATGGGGAAGGCTGGGAGTTCACTCAATGTTGTAATAACAGCCG 664
 QY 184 IleGlyValHisGlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIle 203
 DB 665 CTGGGGAACATGGAGATCCAGTGTGCTGTATGAGTGGNATGAATGTTGCTGGTGC 724
 QY 204 ProLeuSerGluPhe---ValLysGlyIleLysGlnGluValAspAspIle 222
 DB 725 TCTCTGAAGACTCTGCACCCAGATTAGGACTGATAAAGATAAGAACTGGAAGAG 784
 QY 223 ValGlnLysThrLysValAlaGlyIle---GluIleValArgLeuLeuGlyClnGlySer 241
 DB 785 GTTCACAGCAGGTGTTGAGAGTCTCTATGAGTGTCTATGAGTGTATCAAACTCAAAAGGCTACACATCC 844
 QY 242 AlaTyrTyrAlaProGlyAlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAspArg 261
 DB 845 -----TGGGCTATTGGACTCTCTAGCAGATTTGGCAGAGATATTAATGAAGATCTT 898
 QY 262 LysArgValMetValCysSerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHis--- 280
 DB 899 AGCGGGTGCACCCAGTTTCCACCATGATTAAAGGCTCTTACGGGAATAAGGATGATCTC 958
 QY 281 TyrLeuGlyValProCysValIleGlyArgGlyValGluLysIleIleGluLeuGlu 300
 DB 959 TTCCTTAGTGTCTCTGTCATTTGGGACAGAAATGGAATCTCAGACCTTTGTGAAGTGA 1018
 QY 301 LeuThrAlaGlnGluArgGlnGluLeuGlnGlySerIleAspGluValLysGluMetGln 320
 DB 1019 CTGACTCTGAGGAAGAGCCGTTTGAAGAGAGTGCAGATACACTTTGGGGGATCCAA 1078
 QY 321 LysAlaIle 323
 DB 1079 AAGGAGCTG 1087

RESULT 6

US-11-000-688-1061
 ; Sequence 1061, Application US/11000688
 ; Publication No. US20050287544A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BERTUCCI, Francois
 ; APPLICANT: HOULGATTE, Remi
 ; APPLICANT: BIRNBAUM, Daniel
 ; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS
 ; FILE REFERENCE: 1423-R-03
 ; CURRENT APPLICATION NUMBER: US/11/000,688
 ; CURRENT FILING DATE: 2004-12-01
 ; PRIOR APPLICATION NUMBER: US 60/525,987

; PRIOR FILING DATE: 2003-12-01
 ; NUMBER OF SEQ ID NOS: 1596
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 1061
 ; LENGTH: 1661
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial sequences:primer
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)-(1661)
 ; OTHER INFORMATION: lactate dehydrogenase a (LDHA) gene.
 US-11-000-688-1061

Alignment Scores:
 Pred. No.: 3,71e-38 Length: 1661
 Score: 408.00 Matches: 102
 Percent Similarity: 53.6% Conservative: 71
 Best Local Similarity: 31.6% Mismatches: 138
 Query Match: 24.2% Indels: 12
 DB: 12 Gaps: 7

US-09-390-846-2 (1-330) x US-11-000-688-1061 (1-1661)

QY 5 GluLysAsnThrArgProLysIleAlaMetValGlySerGlyMetIleGlyClyThrMet 24
 DB 143 GAACAGACCCCCAGAAATAGATTACAGTCTGTGGGGTGGTGGTGGCATGGCCCTGT 202
 QY 25 AlaPheLeuCysSerLeuArgGluLeuGlyAsp---ValValLeuPheAspValValPro 43
 DB 203 GCCATCATGATATCTAATTAAGGACTTGGCAGACTTGGCAGACTTGGTCTTGTGTGATGTCATCGAA 262
 QY 44 AsnMetProMetGlyLysAlaMetAspIleSerHisAsnSerSerValValAspThrGly 63
 DB 263 GACAAATGAAGGAGAGATGATGATCTCCAACTGACAGCTTTCTTAGAACACCA 322
 QY 64 IleThrValTyrGlySerAsnSerTyrGluCysLeuLysGlyAlaAspValValIleIle 83
 DB 323 AAGATTGCTCTGGCAAGAC---TATAATGTAACCTGCAAACTCCAGCTGTCATATC 379
 QY 84 ThrAlaGlyIleThrLysIleProGlyLysSerAspLysGluTyrSerArgMetAspLeu 103
 DB 380 ACGCTGGGGCAGCTCAGCAAGAGGAGAA-----AGCGCTCTTAATTG 424
 QY 104 LeuProValAsnIleLysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysPro 123
 DB 425 GTCCAGCGTAACTGAACATATTTAAATTCATCTCTTAATGTTGTAATAACAGCCG 484
 QY 124 AsnAlaPheValIleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGlu 143
 DB 485 AACTGCAAGTTGCTTATTGTTTCAATCCAGTGGATATCTTGACCTACGTGGCTTGAAG 544
 QY 144 SerSerGlyLeuProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerArg 163
 DB 545 ATAAGTGTGTTTCCCAAAACCGTGTATTGGAAGTGTGCAATCTGCAATCTGCAATCTG 604
 QY 164 PheArgArgMetIleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetVal 183
 DB 605 TTCCGTTACCTGATGGGGAAGGCTGGGAGTTCACTCAATGTTGTAATAACAGCCG 664
 QY 184 IleGlyValHisGlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIle 203
 DB 665 CTGGGGAACATGGAGATCCAGTGTGCTGTATGAGTGGNATGAATGTTGCTGGTGC 724
 QY 204 ProLeuSerGluPhe---ValLysGlyIleLysGlnGluValAspAspIle 222
 DB 725 TCTCTGAAGACTCTGCACCCAGATTAGGACTGATAAAGATAAGAACTGGAAGAG 784
 QY 223 ValGlnLysThrLysValAlaGlyIle---GluIleValArgLeuLeuGlyClnGlySer 241
 DB 785 GTTCACAGCAGGTGTTGAGAGTGTCTATGAGTGTCTATGAGTGTATCAAACTCAAAAGGCTACACATCC 844

Qy 242 AlaTyrTyrAlaProGlyAlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAspArg 261
 Db 845 -----TGGGCTATTGGACTCTCTGTAGCAGATTGGCAGAGAGTATTAATGAAGAATCTT 898
 Qy 262 LysArgValMetValCysSerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHis--- 280
 Db 899 AGCGGGGTGCACCCNGTTCCACCATGATTAAAGGCTCTTTACGGATTAAGAGATGATGTC 958
 Qy 281 TyrLeuGlyValProCysValIleGlyArgGlyValGluLysIleIleGluLeuGlu 300
 Db 959 TTCCTTAGTGTCTTGTGATTTGGCAGAGATGGAATCTCAGACCTTGTGAAGGTGACT 1018
 Qy 301 LeuThrAlaGlnGluArgGlnGluLeuGlnGlySerIleAspGluValLysGluMetGln 320
 Db 1019 CTGACTTCTGAGGAAGAGCCCGTTTGAAGAAGAGTGCAGATACACACTTTGGGGGATCCAA 1078
 Qy 321 LysAlaIle 323
 Db 1079 AAGGAGCTG 1087

RESULT 7

US-10-955-054A-92
 ; Sequence 92, Application US/10955054A
 ; Publication No. US20050266420A1
 ; GENERAL INFORMATION:
 ; APPLICANT: PUSZTAI, LAJOS
 ; APPLICANT: SYMMANS, W. FRASER
 ; APPLICANT: HESS, KENNETH R.
 ; APPLICANT: AYERS, MARK
 ; APPLICANT: STEC, JAMES
 ; TITLE OF INVENTION: MULTIGENE PREDICTORS OF RESPONSE TO CHEMOTHERAPY
 ; FILE REFERENCE: UTXC:880US
 ; CURRENT APPLICATION NUMBER: US/10/955,054A
 ; CURRENT FILING DATE: 2004-09-30
 ; NUMBER OF SEQ ID NOS: 195
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 92
 ; LENGTH: 1336
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-955-054A-92

Alignment Scores:
 Pred. No.: 4,87e-38 Length: 1336
 Score: 406.00 Matches: 99
 Percent Similarity: 51.5% Conservative: 75
 Best Local Similarity: 29.3% Mismatches: 130
 Query Match: 24.1% Indels: 34
 DB: 8 Gaps: 10

US-09-390-846-2 (1-330) x US-10-955-054A-92 (1-1336)

Qy 5 GluLysAsnThrArgPro-----LysIleAlaMetValGlySerGlyMetIleGlyGly 22
 Db 145 GAAGAGGCAACAGTTCACCAACAAATAGATCACTAGTGGGTGGACAAAGTGTGT--- 201
 Qy 23 ThrMetAlaPheLeuCysSerLeuArgGluLeuGly-----AspValVal 37
 Db 202 -----ATGCGGTGTCTATCAGCATCTCTGGGAAAGTCTCTGCGTGAATGCTGTCT 252
 Qy 38 LeuPheAspValProAsnMetProMetGlyLysAlaMetAspIleSerHisAsnSer 57
 Db 253 CTTGTGGATGTTTGGAGATAAGCTTAAAGGAGAAATATGATGATCTGCAGCATGGGAGC 312
 Qy 58 SerValValAspThrGlyIleThrValTyrGlySerAsnSerTyrGluCysLeuLysGly 77
 Db 313 TTATTTCTCAGACACT---AAAATGTGGCAGATAAAGATTAATCTGTGACCGCCCAAT 369
 Qy 78 AlaAspValValIleIleThrAlaGlyIleThrLysIleProGlyLysSerAspLysGlu 97
 Db 370 TCTAAGATTGTAGTGAATCTGAGGAGTCCGTCAGCAAGAAGGGGAG----- 417
 Qy 98 TrpSerArgMetAspLeuLeuProValAsnIleLysIleMetArgGluValGlyAlaAla 117

Db 418 ---AGTGGCTCAATCTGGTGCAGAGAAAGTTAATCTCTTCAAATTCATTATTCCTCAG 474
 Qy 118 IleLysSerTyrCysProAsnAlaPheValIleAsnIleThrAsnProLeuAspValMet 137
 Db 475 ATCGTCAAGTACAGTCTGATTCATCATAATTTGTGTTTCCCAACCCAGTGGACATTCCT 534
 Qy 138 ValAlaIleGlnGluSerSerGlyLeuProHisIleArgIleCysGlyMetAlaGly 157
 Db 535 ACGTATGTTACCTGGAAACAAAGTGGATTACCCAAACACCCGCTGATTTGGAGTGGATGT 594
 Qy 158 MetLeuAspSerSerArgPheArgMetIleAlaAspLysLeuGluValSerProArg 177
 Db 595 AATCTGATTCGTAGATTTCGTACCTTATGCTGAAACAACTTGGCAATTCATCCCAGC 654
 Qy 178 AspValGlnGlyMetValIleGlyValHisGlyAspHisMetValProLeuSerArgTyr 197
 Db 655 AGCTGCCATGGATGATTTTGGGGGAAACATGGCGACTCAAGTGTGGCTGTGGAGTGGT 714
 Qy 198 AlaThrValAsnGlyIleProLeuSerGluPhe-----Val 209
 Db 715 GTGAATGTGGCAGGTGTTCTCTCCAGGAATTGAATCCAGAAATGGGAACCTGACAATGAT 774
 Qy 210 LysLysGlyTyrIleLysGlnGluValAspAspIleValGlnLysThrLysValAla 229
 Db 775 AGTGAATAATGG-----AAGGAAGTGCATAAGATGGTGGTTGAAAGTGCCTAT--- 822
 Qy 230 GlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaProGlyAlaSer 249
 Db 823 -----GAAGTCATCAAGCTA-----AAAGGATATACCAACTGGGCTATTGGATTAAGT 870
 Qy 250 AlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCysSerCys 269
 Db 871 GTGCTGATCTTATTGAATCCATGTTGAAAAATCTATCCAGGATTCATCCCGTGTCAACA 930
 Qy 270 TyrLeuGlnGlyGlnTyrGlyValGlnAsnHis---TyrLeuGlyValProCysValIle 288
 Db 931 ATGTTAAAGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 990
 Qy 289 GlyGlyArgGlyValGluLysIleGluLeuGluLeuThrAlaGlnGluArgGlnGlu 308
 Db 991 AATCCCGGGGATTAACCCAGCGTTTATCAACCAAGAGCTAAAGGATGATGATGATGATGAT 1050
 Qy 309 LeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaAlaLeu 326
 Db 1051 CTCAAGAAAGTGCAGATATCCCTGTGGGACATCCAGAGGACCTTAAAGACCTG 1104

RESULT 8

US-10-947-249-93
 ; Sequence 93, Application US/10947249
 ; Publication No. US20050287541A1
 ; GENERAL INFORMATION:
 ; APPLICANT: AKIRA NAKAGAWARA
 ; APPLICANT: MIKI OHIRA
 ; APPLICANT: SHIN ISHII
 ; APPLICANT: TAKESHI GOTO
 ; APPLICANT: HIROYUKI KUBO
 ; APPLICANT: TAKAHIRO HIRATA
 ; APPLICANT: YASUKO YOSHIDA
 ; APPLICANT: SAICHI YAMADA
 ; TITLE OF INVENTION: Microarray for Predicting the Prognosis of Neuroblastoma
 ; FILE REFERENCE: 117007
 ; CURRENT APPLICATION NUMBER: US/10/947,249
 ; CURRENT FILING DATE: 2004-09-23
 ; PRIOR APPLICATION NUMBER: US 60/505,614
 ; PRIOR APPLICATION NUMBER: 2003-09-25
 ; NUMBER OF SEQ ID NOS: 200
 ; SOFTWARE: Patent In version 3.1
 ; SEQ ID NO 93
 ; LENGTH: 1336
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens

US-10-947-249-93

Alignment Scores:

Pred. No.: 4,87e-38 Length: 1336
Score: 406.00 Matches: 99
Percent Similarity: 51.5% Conservative: 75
Best Local Similarity: 29.3% Mismatches: 130
Query Match: 24.1% Indels: 34
DB: Gaps: 10

US-09-390-846-2 (1-330) x US-10-947-249-93 (1-1336)

QY 5 GluLysAsnThrArgPro-----LysIleAlaMetValGlySerGlyMetIleGlyGly 22
DB 145 GAAGAGGCAACAGTTCACAAACAATAAGACTACTGTAGTGGGTGGCAAGTTCGT--- 201
QY 23 ThrMetAlaPheLeuCySerLeuArgGluLeuGly-----AspValVal 37
DB 202 -----ATGGCGTGCTATCAGCATCTCGGAAGAAGTCTCTGGCTGATGAATTCGT 252
QY 38 LeuPheAspValProAsnMetProMetGlyLysAlaMetAspIleSerHisAsnSer 57
DB 253 CTTGTGGATGTTTGAAGATAAGCTTAAAGGAGAAATGATGATCTGCAGCATGGAGC 312
QY 58 SerValValAspThrGlyIleThrValTyrGlySerAsnSerTyrGluCysLeuLysGly 77
DB 313 TTAATTTCTCAGACACCT---AAAAATGTGGCAGATAAAGATTAATTTCTGTGACCGCCAAT 369
QY 78 AlaAspValValIleThrAlaGlyIleThrLysIleProGlyLysSerAspLysGlu 97
DB 370 TCTAAGATTGATGTGTAAGTCTCAGAGTCCCTCAGCAAGAGGGAG----- 417
QY 98 TrpSerArgMetAspLeuLeuProValAsnIleLysIleMetArgGluValGlyAlaAla 117
DB 418 ---AGTCGCTCAATCTGTGCAGAGAAATGTTAATGCTTCAAAATTCATTATTCCTCAG 474
QY 118 IleLysSerTyrCyProAsnAlaPheValIleAsnIleThrAsnProLeuAspValMet 137
DB 475 ATCGTCAAGTACAGCTCTGATTCATCATTAATTTGGTTTCCACCCAGTGACATTCCT 534
QY 138 ValAlaAlaLeuGlnGluSerSerGlyLeuProHisArgIleCysGlyMetAlaGly 157
DB 535 ACGTATGTTACTGGAACTAAGTGGATTACCACCAACCGCGTGAATGGAGTGGAGTGT 594
QY 158 MetLeuAspSerSerArgPheArgMetIleAlaAspLysLeuGluValSerProArg 177
DB 595 AATCTGGATTCTGCTAGATTCTGCTACCTATGGCTGAATACTGGCATTCATCCAGC 654
QY 178 AspValGlnGlyMetValIleGlyValHisGlyAspHisMetValProLeuSerArgTyr 197
DB 655 AGCTGCCATGGATGGATTGTTGGGGAACATGCGGACTCAAGTGTGGCTGTGGAGTGGT 714
QY 198 AlaThrValAsnGlyIleProLeuSerGluPhe-----Val 209
DB 715 GTGAATGTGGAGGTGTTCTCTCCAGGAATTTGAATCCAGAAATCGGAACCTGACAATGAT 774
QY 210 LysLysGlyTyrIleLysGlnGluValAspAspIleValGlnLysThrLysValAla 229
DB 775 AGTGAATTTGG-----AAGAGAGTGCATAGATGGTGGTGAAGTCCTAT--- 822
QY 230 GlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaProGlyAlaSer 249
DB 823 -----GAAGTCATCAAGCTA-----AAGGATATACCACTGGGGCTATTGGATTAAGT 870
QY 250 AlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCysSerCys 269
DB 871 GTGGCTGATCTTATTGAATCCATGTGAAATAATCTATCCAGGATTCATCCCGGTGCAACA 930
QY 270 TyrLeuGlnGlyGlnTyrGlyValGlnAsnHis---TyrLeuGlyValProCysValIle 288
DB 931 ATGTTAAGGGGATGATGGCATTCAGAAATGAACTCTTCCTGAGCCTTCATGATATCTC 990
QY 289 GlyGlyArgGlyValGlnLysIleIleGluLeuGluLeuThrAlaGlnGluArgGlnGlu 308

DB 991 AATGCCCGGGGATTAAACAGCGTTATCAACAGAAAGCTAAAGAGTAGATGAGTTCTCTCAG 1050
QY 309 LeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaAlaLeu 326
DB 1051 CTCAAGAAAGTGCAGATACCTCTGGGACATCCAGAGGACCTAAAGACCTG 1104
RESULT 9
US-10-506-443A-1
; Sequence 1, Application US/10506443A
; Publication No. US20060013817A1
; GENERAL INFORMATION:
; APPLICANT: Sahin Dr., Ugur
; APPLICANT: Tureci Dr., Ozlem
; APPLICANT: Koglowski Dr., Michael
; TITLE OF INVENTION: Genetic Products Differentially Expressed in Tumors and Use There
; FILE REFERENCE: 342-3PCT
; CURRENT APPLICATION NUMBER: US/10/506,443A
; CURRENT FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1171
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-506-443A-1
Alignment Scores:
Pred. No.: 5,43e-38 Length: 1171
Score: 405.00 Matches: 101
Percent Similarity: 52.9% Conservative: 72
Best Local Similarity: 30.9% Mismatches: 130
Query Match: 24.0% Indels: 24
DB: Gaps: 7
US-09-390-846-2 (1-330) x US-10-506-443A-1 (1-1171)
QY 5 GluLysAsnThrArgProLysIleAlaMetValGlySerGlyMetIleGlyGlyThrMet 24
DB 104 GATGAAACTCCAGTGTAAATATTACTATTCTTGGAACTGGTGGCTAGGCATGGCTTGT 163
QY 25 AlaPheLeuCySerSerLeuArgGluLeuGlyAsp---ValValLeuPheAspValValPro 43
DB 164 GCTATTAGTATCTTACTGAAGGATTTGGCTGTGAACCTTGGCTTTGTTGATGTTGCATTG 223
QY 44 AsnMetProMetGlyLysAlaMetAspIleSerHisAsnSerSerValValAspThrGly 63
DB 224 GACAACTGAAGGAGAAATGATGATCTTCAGCATGGCAGTCTTTCTTTAGTACTTCA 283
QY 64 IleThrValTyrGlySerAsnSerTyrGluCysLeuLysGlyAlaAspValValIleIle 83
DB 284 AAGATTACTTCTGGAAAGAT---TACAGTGTATTCGAAACTCCAGATAGTATTGCTC 340
QY 84 ThrAlaGlyIleThrLysIleProGlyLysSerAspLysGluTyrSerArgMetAspLeu 103
DB 341 ACAGCAGGTGCAAGCAGCAGGAGGAGAA-----ACTCGCTTGGCCCTG 385
QY 104 LeuProValAsnIleLysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysPro 123
DB 386 GTCCAACTGAATGGCTTATTAATCAATCAATCTCTGCGCATAGTCCATTATAGTCTCT 445
QY 124 AsnAlaPheValIleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGlu 143
DB 446 GATTGTAATAATCTTGTGTTTCAATCCAGTGGATATTTGACATATATAGTCTGGAG 505
QY 144 SerSerGlyLeuProHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArg 163
DB 506 ATAAGTGGCTTACCTGTAACCTCGTGAATTTGGAAGTGGTGTGTAATCTAGACTCTGCCCGT 565
QY 164 PheArgMetIleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetVal 183
DB 566 TTCCGTTACCTAATTTGGAGAAAGTTGGTGTGCCACCCCAAGCTGCCATGGTGGATT 625

Qy 184 IleGlyValHisGlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIle 203
Db 626 ATTGAGAACATGGTGAATCTAGTGTGCCCTTATGGAGTGGGTGAATGTTCTGGTGT 685
Qy 204 ProLeu-----SerGluPheValLysGlyTyrIleLys 215
Db 686 GCTCTGAAGACTCTGACCCCTAAATTTAGGAACGGATTTCAGATAAGAACACTGG----- 739
Qy 216 GlnGluValAspAspIleValGlnLysThrLysValAlaGlyGlyIleValArg 235
Db 740 -----AAAAATATCATCAAAACAAAGTATTCAAAAGTCCCTATGAATATATCAAG 787
Qy 236 LeuLeuGlyGlnGlySerAlaTyrTyrAlaProGlyAlaSerAlaIleGlnMetAlaGlu 255
Db 788 CTGAAGGGGTATACCTCT-----TGGGCTATTGGACTCTCTGATGGATCTGTAGGA 841
Qy 256 SerTyrLeuLysAspArgLysArgValMetValCysSerCysTyrLeuGlnGlyGlnTyr 275
Db 842 TCCATTTTCAAAAATCTTAGGAGAGTGCACCCAGTTTCCACCATGGTTAAGGGATTATAT 901
Qy 276 GlyValGlnAsnHis---TyrLeuGlyValProCysValIleGlyCysArgGlyValGlu 294
Db 902 GGAATAAAGAAAGAACTCTTTCTCAGTATCCCTTGTGTCTTGGGGGGAATGGTGTCTCA 961
Qy 295 LysIleIleGluLeuGluLeuThrAlaGlnGluA:GlnGluLeuGlnGlySerIleAsp 314
Db 962 GATGTTGTGAAAATTAACCTTGAATCTCTGAGGAGGAGCCCTTTTCAAGNAGAGTGCAGAA 1021
Qy 315 GluValLysGluMetGlnLys 321
Db 1022 ACATTTTGGAAATATTCAAAAG 1042

RESULT 10

US-10-793-626-1411
; Sequence 2383, Application US/10793626
; Publication No. US20050255478A1

GENERAL INFORMATION:

; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1411
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence

US-10-793-626-1411

Alignment Scores:
Pred. No.: 1,12e-37 Length: 969
Score: 401.50 Matches: 93
Percent Similarity: 53.0% Conservative: 75
Best Local Similarity: 29.3% Mismatches: 140
Query Match: 23.8% Indels: 9
DB: 8 Gaps: 4

US-09-390-846-2 (1-330) x US-10-793-626-1411 (1-969)

Qy 11 LysIleAlaMetValGlySerGlyMetIleGlyGlyThrMetAlaPheLeuCysSerLeu 30
Db 37 AAAGTGTGTTAGTAGGACCGTCCGTAGGTCAAGTATGCAATTCCTATGCTGACT 96
Qy 31 ArgGluLeuGlyAsp---ValValLeuPheAspValProAsnMetProMetGlyLys 49
Db 97 CAAGGAATTCAGATGAATTTGTAAATTATTTGATATTTGCAAAAGATAAAGTGAAGCAGAC 156

Qy 50 AlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyrGlySer 69
Db 157 GTTAAAGATTTAAACCATGGTGCACTTTACAGTTCTTCACCAGTACGTGTA---AAAGCT 213
Qy 70 AsnSerTyrGluCysLeuLysGlyAlaAspValValIleIleThrAlaGlyIleThrLys 89
Db 214 GGAGAAATATGAAGATTGTAAAGATGCAGATTAGTTAGTTATTATACAGCAGGTGCACCTCAA 273
Qy 90 IleProGlyLysSerAspLysGluTyrSerArgMetAspLeuLeuProValAsnIleLys 109
Db 274 AAACCGGGTGA-----ACTCGTTTCAACACTTGTGTGAAATAACTACTAAA 318
Qy 110 IleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAlaPheValIleAsn 129
Db 319 ATCATGAAAAGTATCGTAACCTAGTGTGCATGATAGTGGCTTTGTGTTCTTCTCTAATT 378
Qy 130 IleThrAsnProLeuAspValMetValAlaAlaGlnGlnGlnGlnGlnGlnGlnGln 149
Db 379 GCTGCAAAACCCAGTGTATCTTAACACGTTATGTTAAAGAGATTACAGGTTTACCAGCT 438
Qy 150 HisArgIleCysGlyMetAlaGlyMetLeuAspSerArgPheArgMetIleAla 169
Db 439 GAACGTGTTATTGGTCTCTGACAGTGTCTGATAGTCAAGATTTCAGATATTTAATAAGT 498
Qy 170 AspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHisGlyAsp 189
Db 499 AAAGAAATAGGTGTTACATCAAGTAGTGTTCACCGTAGCATTTATAGTGAACATGGTGAC 558
Qy 190 HisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGluPheVal 209
Db 559 TCTGAACCTTGCAGTTTGGTCTCAAGCAACCGTTGGAGGTATTTTCAGTGTATGATACATTG 618
Qy 210 LysLysGlyTyrIleLysGlnGluValAspIleValGlnLysThrLysValAla 229
Db 619 AAAGAAAGAACTGGTAGCGATGCTAAAGCGAATGAAATTTATATTAATACAGAGATGCT 678
Qy 230 GlyGlyGluIleValArgLeuLeuGlnGlySerAlaTyrTyrAlaProGlyAlaSer 249
Db 679 GCTTACGATATCATTTCAA-----GCTAAAGGATCTAGCTATTATTTAGTATAGCTTAGCA 732
Qy 250 AlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCysSerCys 269
Db 733 CTATTACGTATTTCTAAAGCTTTACTAAATAATGAAATAGTATTTTTCACACAGTTTCTAGT 792
Qy 270 TyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysValIleGly 289
Db 793 CAACCTTAATGTCATATGGAATTTAACGATGTTTATCTTGGCTTACCACACTTATCAAT 852
Qy 290 GlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGluArgGlnGluLeu 309
Db 853 CAABAATGGTGCAGTTTAAATTTATGAAACACCATTAATGATGATGAACTACAATTACTTA 912
Qy 310 GlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaAlaLeu 326
Db 913 GAAAAATCAGTGAAAACTTTAGAGACACTTTATGATCTTATATAAAACATTTTA 963

RESULT 11

US-10-793-626-2383
; Sequence 2383, Application US/10793626
; Publication No. US20050255478A1

GENERAL INFORMATION:

; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2383
; LENGTH: 969
; TYPE: DNA

		1329	AAAGAATTAGGTGTTACATCAAGTAGTGTTCACGCTAGCATATTAGGTGAACATGGTGC	1270
Db				
		190	HiSmetValProLeuSerArgTyTyrAlaThrValAsnGlyIleProLeuSerGluPheVal	209
Qy				
		1269	TCTGAACCTTCGAGTTGGTCTCAAGCAAAAGCTTGGAGGTATTTTCAGTGTATGAATACATGGT	1210
Db				
		210	LysLysGlyTrpIleLysGlnGluValAspAspileValGlnLysThrLysValAla	229
Qy				
		1209	ANAGAAGAAACTGGTAGCGATGCTAAAGCGAATGAATTTATTAATACAGAGATGCT	1150
Db				
		230	GlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyTyrAlaProGlyAlaSer	249
Qy				
		1149	GCITTACGATATCAITTCAA-----GCTAAAGGATCTACGTATTATGTGTATAGCTCTAGCA	1096
Db				
		250	AlaIleGlnMetAlaGluSerTyLeuLysAspArgLysArgValMetValCysSerCys	269
Qy				
		1095	CTATTCGTATTTCTAAAGCTTTTACTAAATAATAAGAAATAGTATTGTGACAGTTCCTAGT	1036
Db				
		270	TyrLeuGlnGlyGlnTyGlyValGlnAsnHisTyLeuGlnGlyValProCysValIleGly	289
Qy				
		1035	CAACTTAATGGTCAATATGGATTTTAACGATGTTTATCTTGGCTTTACCACACATTTATCAAT	976
Db				
		290	GlyArgGlyValGluLysIleIleGluLeuLeuThrAlaGlnGluArgGlnGluLeu	309
Qy				
		975	CAAAATCGTCAGTTAAATTTATGAACACCATTTAAATGATACGAACTACAAATTACTA	916
Db				
		310	GlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaLeu	326
Qy				
		915	GAATAAATCAGTGAATACTTTTGAAGACACATTTATGATCTCTATAAACATTTA	865
Db				

RESULT 13

```

RES001 13
US-10-793-626-3793
; Sequence 3793, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3793
; LENGTH: 4093
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3793

```

Alignment Scores:	6.98e-37	Length:	4093
Pred. No.:		Matches:	93
Score:	401.50	Conservative:	75
Percent Similarity:	53.0%	Mismatches:	140
Best Local Similarity:	29.3%	Indels:	9
Query Match:	23.8%	Gaps:	4
DB:	8		

US-09-390-846-2 (1-330) x US-10-793-626-3793 (1-4093)

Qy	11	LysIleAlaMetValGlySerGlyWetIleCglyThrMetAlaPheLeuCysSerLeu	30
Db	1105	AAAGTGTATTTAGTAGGACGGTTCCTAGGTTCAAGATTATGCATTTCGTATCGTGACT	1164
Qy	31	ArgGluLeuGlyAsp---ValValLeuPheaspValValProAenMetProMetGlyLys	49
Db	1165	CAAGGAATTGCAGATGAATTTGTAAATTATTCATATTGCCAAAGAATAAAGTGGAGACAGAC	1224
Qy	50	AlaMetAspIleSerHisAasnSerValValAspThrGlyIleThrValTyrglySer	69

; PRIOR APPLICATION NUMBER: US 60/570,425
 ; PRIOR FILING DATE: 2004-05-11
 ; NUMBER OF SEQ ID NOS: 7285
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 676
 ; LENGTH: 1317
 ; TYPE: DNA
 ; ORGANISM: Cricetinae gen. sp.
 ; US-11-128-061-676

Alignment Scores:
 Pred. No.: 7,48e-37 Length: 1317
 Score: 396.00 Matches: 102
 Percent Similarity: 52.3% Conservativeness: 69
 Best Local Similarity: 31.2% Mismatches: 132
 Query Match: 23.5% Indels: 24
 DB: 12 Gaps: 7

US-09-390-846-2 (1-330) x US-11-128-061-676 (1-1317)

QY	5	GlulysAsnThrArgProLysAlaMetValGlySerGlyMetIleGlyThrMet	24
DB	92	GATAAAACCTCCGGTGAAGATCACTGTGGGAGTTGGGATGTCGGCATCGCGTGT	151
QY	25	AlaPheLeuCysSerLeuArgGluLeuGlyAsp---ValValLeuPheAspValValPro	43
DB	152	GCTATTAGTATCTTACTGAAGGATTTGGCCGATGAACCTGGCCCTTGTGACACTGATACC	211
QY	44	AsnMetProMetGlyLysAlaMetAspIleSerHisAsnSerValValAspThrGly	63
DB	212	GACAACTGAAGGAGAGACGCTGGATCTTCTGCGGAGTCTTTTCTTCAACACTCA	271
QY	64	IleThrValTyrGlySerAsnSerTyrGluCysLeuLysGlyAlaAspValValIle	83
DB	272	AAGATTGTCTCGAAAAGAT---TACAGTGTGTCTGCCAACTCCAAATTTAGTTATCATC	328
QY	84	ThrAlaGlyIleThrLysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeu	103
DB	329	ACAGCCGGTCTAGGCAAGAGTGGGGAA-----ACAGCCCTTGACCTG	373
QY	104	LeuProValAsnIleLysIleMetArgGluValGlyAlaIleLysSerTyrCysPro	123
DB	374	GTCACGGTAAGCTGTCTATCATGAATCCATCGTCTCGCTTAGTGCAACATAGTCCT	433
QY	124	AsnAlaPheValIleAsnIleThrAsnProLeuAspValMetValAlaLeuGlnGlu	143
DB	434	GACTGTAAATGCTGATCTCTCAACCCAGTGGATATTTGACTTATGTGCTGGAG	493
QY	144	SerSerGlyLeuProHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArg	163
DB	494	ATAAGTGGCTGCTGCAACTCGTGTAAATTTGGAAGTGGCTGTAACTTAGATTCTGCCCGT	553
QY	164	PheArgArgMetIleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetVal	183
DB	554	TTCCGTTACTTATTTGGAGAGAAGTTGGGTGTCACCTTACAACTGTCGCGTGGGT	613
QY	184	IleGlyValHisGlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyLe	203
DB	614	CTTGGAGACACGGTGACTCTAGTGTGCCCTTATGGAGTGGTGTGAACGTTGTCGGCGTA	673
QY	204	ProLeu-----SerGluPheValLysLysGlyTyrIleLys	215
DB	674	CTTCTGAAGTCACTGAACCCAGCTTAGGAGCGACTCAGACCAAGAGCACTGGAAAAG	733
QY	216	GlnGluValAspAspIleValGlnLysThrLysValAlaGlyGlyGluIleValArg	235
DB	734	-----ATCCACAACCAAGTGGTGGAGAGTGGCTATGAGGTCCTCCCG	775
QY	236	LeuLeuGlyGlnGlySerAlaTyrAlaProGlyAlaSerAlaIleGlnMetAlaGlu	255
DB	776	CTGAAGGGCTATACATCT-----TGGGCTATTGGGCTGTCCTGGACAGATCTGGCAAGA	829
QY	256	SerTyrLeuLysAspArgLysArgValMetValCysSerCysTyrTyrLeuGlnGlyInTyr	275

; PRIOR APPLICATION NUMBER: US 60/570,425
 ; PRIOR FILING DATE: 2004-05-11
 ; NUMBER OF SEQ ID NOS: 7285
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 676
 ; LENGTH: 1317
 ; TYPE: DNA
 ; ORGANISM: Cricetinae gen. sp.
 ; US-11-128-061-676

Alignment Scores:
 Pred. No.: 7,48e-37 Length: 1317
 Score: 396.00 Matches: 102
 Percent Similarity: 52.3% Conservativeness: 69
 Best Local Similarity: 31.2% Mismatches: 132
 Query Match: 23.5% Indels: 24
 DB: 12 Gaps: 7

US-11-128-049-676

DB	830	TCAATCTCTGAAGAATCTTAGGAGGGTGCATCCCTTTCCACACTGGTTAAGGCTTATAT	899
QY	276	GlyValGlnAsnHis---TyrLeuGlyValProCysValIleGlyGlyArgGlyValGlu	294
DB	890	GAGATAAAGGAGAAATCTTCTCAGCATTCCTTGTATCTTTGGGACAAAACGGCATCACA	949
QY	295	LysIleIleGluLeuLeuThrAlaGlnGluArgGlnGluLeuGlnGlySerIleAsp	314
DB	950	GACATTTGTGAAGTGAACCTGAATCCCGAGGAGGAGGCCCTTCTCAGGAAGAGTGGCGAA	1009
QY	315	GluValLysGluMetGlnLys	321
DB	1010	ACACTCTGGAACGTCCTCAAAAG	1030

RESULT 15
 US-11-128-049-676
 ; Sequence 676, Application US/11128049
 ; Publication No. US20060010513A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Melville, Mark W.
 ; APPLICANT: Charlebois, Timothy S.
 ; APPLICANT: Mounts, William M.
 ; APPLICANT: Hann, Louane E.
 ; APPLICANT: Sinacore, Martin S.
 ; APPLICANT: Leonard, Mark W.
 ; APPLICANT: Brown, Eugene L.
 ; APPLICANT: Miller, Christopher P.
 ; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND MET
 ; FILE REFERENCE: 01997.027700
 ; CURRENT APPLICATION NUMBER: US/11/128,049
 ; PRIOR FILING DATE: 2005-05-11
 ; PRIOR FILING DATE: 2004-05-11
 ; NUMBER OF SEQ ID NOS: 7285
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 676
 ; LENGTH: 1317
 ; TYPE: DNA
 ; ORGANISM: Cricetinae gen. sp.
 ; US-11-128-049-676

Alignment Scores:
 Pred. No.: 7,48e-37 Length: 1317
 Score: 396.00 Matches: 102
 Percent Similarity: 52.3% Conservativeness: 69
 Best Local Similarity: 31.2% Mismatches: 132
 Query Match: 23.5% Indels: 24
 DB: 12 Gaps: 7

US-09-390-846-2 (1-330) x US-11-128-049-676 (1-1317)

QY	5	GlulysAsnThrArgProLysAlaMetValGlySerGlyMetIleGlyThrMet	24
DB	92	GATAAAACCTCCGGTGAAGATCACTGTGGGAGTTGGGATGTCGGCATCGCGTGT	151
QY	25	AlaPheLeuCysSerLeuArgGluLeuGlyAsp---ValValLeuPheAspValValPro	43
DB	152	GCTATTAGTATCTTACTGAAGGATTTGGCCGATGAACCTGGCCCTTGTGACACTGATACC	211
QY	44	AsnMetProMetGlyLysAlaMetAspIleSerHisAsnSerValValAspThrGly	63
DB	212	GACAACTGAAGGAGAGACGCTGGATCTTCTGCGGAGTCTTTTCTTCAACACTCA	271
QY	64	IleThrValTyrGlySerAsnSerTyrGluCysLeuLysGlyAlaAspValValIle	83
DB	272	AAGATTGTCTCGAAAAGAT---	

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 2, 2006, 23:58:53 ; Search time 207 Seconds
(without alignments)
2833.794 Million cell updates/sec

Title: US-09-390-846-2

Perfect score: 1688
Sequence: 1 MAVENKTRPIAMVSGMIGSIDEVKEMQKAIADLASK 330

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abss/ABSSWEB pool/US09390846/runat 02032006 104236 4927/app query.fasta_1
-DB=Issued Patents NA -QWTF=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -LOCALALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss03p -USER=US09390846 @CGN 1 1 290 @runat 02032006 104236 4927
-NCPU=6 -ICPU=3 -NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOF=6
-FGAPEXT=7 -YGAPEXT=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

1: /cgn2_6/ptodata/1/ina/1 COMB.seq:
2: /cgn2_6/ptodata/1/ina/5 COMB.seq:
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:
5: /cgn2_6/ptodata/1/ina/H COMB.seq:
6: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq:
7: /cgn2_6/ptodata/1/ina/PP COMB.seq:
8: /cgn2_6/ptodata/1/ina/RE COMB.seq:
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1688	100.0	1679	3	US-08-676-882-1
2	1034.5	61.3	1785	3	US-09-216-393B-311
3	1034.5	61.3	1785	3	US-09-216-393B-313
4	831.5	49.3	951	6	PCT-US94-03796-1
5	706	41.8	1912	2	US-08-270-013B-1
6	706	41.8	1912	2	US-08-838-418-1
7	694	41.1	870	3	US-09-902-540-6244
8	694	41.1	1634	3	US-09-902-540-341
9	641	38.0	978	3	US-09-134-001C-2696

10	632	37.4	5024	2	US-08-920-812-7	Sequence 7, Appli
11	632	37.4	5024	2	US-08-920-827-7	Sequence 7, Appli
12	632	37.4	5024	2	US-08-921-177-7	Sequence 7, Appli
13	632	37.4	5024	2	US-08-362-577C-7	Sequence 7, Appli
14	632	37.4	5024	2	US-08-920-828-7	Sequence 7, Appli
15	464	27.5	975	3	US-09-107-532A-2832	Sequence 2832, Ap
16	434.5	25.7	580073	3	US-08-545-528D-1	Sequence 1, Appli
17	426	25.2	957	3	US-09-134-000C-1012	Sequence 1012, Ap
18	408	24.2	1661	3	US-09-166-350-31	Sequence 31, Appli
19	408	24.2	1661	3	US-09-220-132-36	Sequence 36, Appli
20	408	24.2	1661	3	US-09-023-655-1340	Sequence 1340, Ap
21	406	24.1	1260	3	US-09-166-350-22	Sequence 22, Appli
22	406	24.1	1383	3	US-09-949-016-2131	Sequence 2131, Ap
23	405	24.0	1173	3	US-09-949-016-5381	Sequence 5381, Ap
24	405	24.0	1173	3	US-09-949-016-5382	Sequence 5382, Ap
25	401.5	23.8	969	3	US-09-710-279-1411	Sequence 1411, Ap
26	401.5	23.8	969	3	US-09-710-279-2383	Sequence 2383, Ap
27	401.5	23.8	993	3	US-09-134-001C-1491	Sequence 1491, Ap
28	401.5	23.8	3019	3	US-09-710-279-4355	Sequence 4355, Ap
29	401.5	23.8	4093	3	US-09-710-279-3793	Sequence 3793, Ap
30	396.5	23.5	14333	3	US-08-956-171E-57	Sequence 57, Appli
31	396.5	23.5	14333	3	US-08-781-986A-57	Sequence 57, Appli
32	395.5	23.4	924	3	US-09-107-532A-620	Sequence 620, App
33	391	23.2	1240	2	US-08-869-506-1	Sequence 1, Appli
34	391	23.2	1240	2	US-09-128-967-1	Sequence 1, Appli
35	388	23.0	995	3	US-09-107-433-2149	Sequence 2149, Ap
36	388	23.0	995	3	US-09-583-110-1930	Sequence 1930, Ap
37	388	23.0	5278	3	US-08-961-527-227	Sequence 227, App
38	385.5	22.8	3264	3	US-08-956-171E-424	Sequence 424, App
39	385.5	22.8	3264	3	US-08-781-986A-424	Sequence 424, App
40	377	22.3	950	3	US-09-830-230A-563	Sequence 563, App
41	375	22.2	5449	3	US-09-546-990-1	Sequence 1, Appli
42	374	22.2	240157	3	US-09-949-016-16264	Sequence 16264, A
43	371	22.0	1146	3	US-09-711-681-1	Sequence 1, Appli
44	371	22.0	1146	3	US-10-274-266-1	Sequence 1, Appli
45	371	22.0	3144	3	US-09-711-681-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-08-676-882-1
; Patent No. 6100241
; GENERAL INFORMATION:
; APPLICANT: Kok, Jacobus Johannes
; APPLICANT: van den Boogaart, Paul
; APPLICANT: Vermeulen, Arnoldus Nicolaas
; TITLE OF INVENTION: Coccidiosis poultry vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Akzo No. 6100241el Patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/676,882
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 258-5200
; TELEFAX: (301) 977-0847
; INFORMATION FOR SEQ ID NO: 1:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1679 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Eimeria acervulina
; DEVELOPMENTAL STAGE: Schizont
; IMMEDIATE SOURCE:
; CLONE: EASC2_1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 280..1269
; OTHER INFORMATION: /function= "Eimeria lactate"
; OTHER INFORMATION: dehydrogenase"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..51
; OTHER INFORMATION: /label= pBluescriptII
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1624..1679
; OTHER INFORMATION: /label= pBluescriptII
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 45..54
; OTHER INFORMATION: /label= EcoRI-linker
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1621..1630
; OTHER INFORMATION: /label= EcoRI-linker
;
US-08-676-882-1
Alignment Scores:
Pred. No.: 4,41e-197 Length: 1679
Score: 1688.00 Matches: 330
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-09-390-846-2 (1-330) x US-08-676-882-1 (1-1679)
QY 1 MetAlaValPheGluLeuAsnThrArgProLysIleAlaMetValGlySerGlyMetIle 20
DB 280 ATGGCGGCTTCGAGAAGAATACACGCCCAAGATTGCTATGGTGGGCTCCGGTATGATT 339
QY 21 GlyGlyThrMetAlaPheLeuCysSerLeuArgGluLeuGlyAspValValLeuPheAsp 40
DB 340 GGAGGCACCATGGCTTTCCTGTGCAGCTTGAGGGAACCTCGGAGATGTTGTCTCTTCGAC 399
QY 41 ValValProAsnMetProMetGlyLysAlaMetAspIleSerHisAsnSerSerValVal 60
DB 400 GTGTACCGAATCCCGATGGGGAAGCGATGGATATATCGCAAAATTCGTGCGTGGTT 459
QY 61 AspThrGlyIleThrValTyrGlySerAsnSerTyrGluCysLeuLysGlyValAlaAspVal 80
DB 460 GACACGGGTATPAACAGTATACGGCTCAATTCATACAGATGCTTGAAGGGTGGGACGTA 519
QY 81 ValIleIleThrAlaGlyIleThrLysIleProGlyLysSerAspLysGluTrpSerArg 100
DB 520 GTAATAATAACGACGGGATAACAAGATACCCGGAAAGCGATAAAGAATGGTCTAGA 579
QY 101 MetAspLeuLeuProValAsnIleLysIleMetArgGluValGlyValAlaIleLysSer 120
DB 580 ATGGATCTATTACCTGTGAATATAAGATATATAGGAGGGTGGGTCGACGAATTAATATCT 639
QY 121 TyrCysProAsnAlaPheValIleAsnIleThrAsnProLeuAspValMetValAlaAla 140
DB 640 TACTGTCTTAATGCAATTTGTTATTATTAACAATCTTTTAGATGTGATGGTAGCTGCT 699

```

```

QY 141 LeuGlnGluSerSerGlyLeuProHisHisArgIleCysGlyMetAlaGlyMetLeuAsp 160
DB 700 CTTCAAGAGTCATCAGGACTACCTCATCATAGAAATCTGCGGTATGGCTGGGATGCTGAT 759
QY 161 SerSerArgPheArgArgMetIleAlaAspLysLeuGluValSerProArgAspValGln 180
DB 760 AGCTCTCGTTTTAGACGTATGATAGCTGATAAATTAGAAGTCTCTCTCCTAGAGATGTACAG 819
QY 181 GlyMetValIleGlyValHisGlyAspHisMetValProLeuSerArgTyrAlaThrVal 200
DB 820 GGGATGTCATAGTGTACACGGCGCATCATATGTGCCCCCTAAAGTAGATATGCAACAGTT 879
QY 201 AsnGlyIleProLeuSerGluPheValLysValLysGlyTyrIleLysGlnGluValAsp 220
DB 880 ACGGCATCCCGCTTCTGAGTTTGTTAAGAGGGCTGGATCAAGCAAGAAAGATAGAT 939
QY 221 AspIleValGlnLysThrLysValAlaGlyGlyLysIleValArgLeuLeuGlyGlnGly 240
DB 940 GATATCGTTCAGAAAGCAAGGTCGCTGGAGGAGAGATCGTACGCCCTATTAGGACAAGGC 999
QY 241 SerAlaTyrTyrAlaProGlyAlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAsp 260
DB 1000 TCTGCTTACTATGCTCCAGGGGCTTCAGCTATTTCAGATGGCTGAGAGCTATCTAAAGGAT 1059
QY 261 ArgLysArgValMetValCysSerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHis 280
DB 1060 AGAAAGAGAGTGTGTTGCTCTTGCTACTTGAAGGACAAATATGGTGTACAGATCAC 1119
QY 281 TyrLeuGlyValProCysValIleGlyGlyArgGlyValGluLysIleIleGluLeuGlu 300
DB 1120 TACTTAGAGTACCTTGTGTATCGGTGGGAGAGGTGTTGAGAAGATTATTAGTTAGAA 1179
QY 301 LeuThrAlaGlnGluArgGlnGluLeuGlnGlySerIleAspGluValLysGluMetGln 320
DB 1180 TTGACCGCACAAAGAACAGCAGGAGCTTCAGGGATCTATCGATGAGGTTAAGGAGATCGAG 1239
QY 321 LysAlaIleAlaAlaLeuAspAlaSerLys 330
DB 1240 AAGGCTATTGCTGCTTTGATGATCAATCAAG 1269

RESULT 2
US-09-216-393B-311
; Sequence 311, Application US/09216393B
; Patent No. 6514694
; GENERAL INFORMATION:
; APPLICANT: Milhausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393B
; PRIOR FILING DATE: 1998-12-18
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 311
; LENGTH: 1785
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(75)
; OTHER INFORMATION:
;
US-09-216-393B-311
Alignment Scores:
Pred. No.: 1.05e-116 Length: 1785
Score: 1034.50 Matches: 193
Percent Similarity: 77.3% Conservative: 55
Best Local Similarity: 60.1% Mismatches: 72
Query Match: 61.3% Indels: 1
DB: 3 Gaps: 1

```

US-09-390-846-2 (1-330) x US-09-216-393B-311 (1-1785)

```

QY 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
DB 109 AGAAGAAGTGGCCATGATGGCTCTGGCATGATGGTGACATATGGCTACCTGTGC 168
QY 29 SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMetGly 48
DB 169 GCTCTCCGTGACCTGCTGACCTGCTCTCTACGATGTTGTCAGAGGTATGCCAGGGT 228
QY 49 LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyrGly 68
DB 229 AAGGCTCTGACCTGAGCCATGACCTCCGGTGGACACCAACGTTTCGGTCCGTGCT 288
QY 69 SerAsnSerTyrGlu---CysLeuLysGlyAlaAspValValIleThrAlaGlyIle 87
DB 289 GAGTACTCTTACGAGCGCGCTACCGTGGGACCTGCTGATGTCGATGTCGCTGCTG 348
QY 88 ThrLysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProValAsn 107
DB 349 ACCAAGGTGGCGGACCGGACCTCCGAGTGGAGCCGAACGATCTGCTCCCGTTCAAC 408
QY 108 IleLysIleMetArgGluValGlyAlaIleLysSerTyrCysProAsnAlaPheVal 127
DB 409 TCGAAGATCATTCGCGAGATCGTCAGAACATCAAGAAGTACTGCCCAAGACCTTCATC 468
QY 128 IleAsnIleThrAsnProLeuAspValMetValAlaLeuGlnGluSerSerGlyLeu 147
DB 469 ATCGTGTCTACCAACCGCTGCTGATGTCGATGTCGAGGTCCTGCGGAGCTCTGCGGTC 528
QY 148 ProHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMet 167
DB 529 CCGACCAACATGATCTCGGTATGCGCTGATGCTGACCTCTGCTGCTGCTGCTGCTGCTG 228
QY 168 IleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHis 187
DB 589 GTCCCGACGCGCTGCTGCTCTCCCGGACGCTGCGAGCCACCGCTCATCGGAGGAGGAG 248
QY 188 GlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGlu 207
DB 649 GCGGACTGCATGCTCCGCTTGTCCGCTACATTACCGTGAACGACTACCGGAGGAGGAG 267
QY 207 ValAlaGlyGlyValGluLeuValArgLeuGlyGlnGlySerAlaTyrTyrAlaProGly 247
DB 709 TTCTACCAAGGACGCGGTAGTCACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 268
QY 248 AlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCys 267
DB 829 GCATCCGCTGTCGCCATGCGCAACATCTTCTTGAACGACGAGAAAGCGGCTCATCCGCTG 888
QY 268 SerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysVal 287
DB 889 AGTGCTGACTGCACCGAGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 948
QY 288 IleGlyGlyArgGlyValGlyLysIleGluLeuGluLeuThrAlaGlnGluArgGln 307
DB 949 ATTGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1008
QY 308 GluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaLeuAsp 327
DB 1009 CAGTTCAGAGAGTCCGTCGACGAGCTCATCGGCTCAACAGGCGGTTGCTGCTTCAG 1068
QY 328 Ala 328
DB 1069 GCG 1071

```

RESULT 3

US-09-216-393B-313

; Sequence 313, Application US/09216393B

Patent No. 6514694

GENERAL INFORMATION:

APPLICANT: Milhausen, Michael James

TITLE OF INVENTION: TOPOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF

FILE REFERENCE: TX-1-C2

CURRENT APPLICATION NUMBER: US/09/216,393B

CURRENT FILING DATE: 1998-12-18

PRIOR APPLICATION NUMBER: 08/994,825

PRIOR FILING DATE: 1997-12-19

NUMBER OF SEQ ID NOS: 366

SOFTWARE: PatentIn version 3.1

SEQ ID NO 313

LENGTH: 1785

TYPE: DNA

ORGANISM: Toxoplasma gondii

US-09-216-393B-313

Alignment Scores:

```

Pred. No.: 1,05e-116 Length: 1785
Score: 1034.50 Matches: 193
Percent Similarity: 77.3% Conservativity: 55
Best Local Similarity: 60.1% Mismatches: 72
Query Match: 61.3% Indels: 1
DB: 3 Gaps: 1

```

US-09-390-846-2 (1-330) x US-09-216-393B-313 (1-1785)

```

QY 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
DB 109 AGAAGAAGTGGCCATGATGGCTCTGGCATGATGGTGACATATGGCTACCTGTGC 168
QY 29 SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMetGly 48
DB 169 GCTCTCCGTGACCTGCTGACCTGCTCTCTACGATGTTGTCAGAGGTATGCCAGGGT 228
QY 49 LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyrGly 68
DB 229 AAGGCTCTGACCTGAGCCATGACCTCCGGTGGACACCAACGTTTCGGTCCGTGCT 288
QY 69 SerAsnSerTyrGlu---CysLeuLysGlyAlaAspValValIleThrAlaGlyIle 87
DB 289 GAGTACTCTTACGAGCGCGCTACCGTGGGACCTGCTGATGTCGATGTCGCTGCTG 348
QY 88 ThrLysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProValAsn 107
DB 349 ACCAAGGTGGCGGACCGGACCTCCGAGTGGAGCCGAACGATCTGCTCCCGTTCAAC 408
QY 108 IleLysIleMetArgGluValGlyAlaIleLysSerTyrCysProAsnAlaPheVal 127
DB 409 TCGAAGATCATTCGCGAGATCGTCAGAACATCAAGAAGTACTGCCCAAGACCTTCATC 468
QY 128 IleAsnIleThrAsnProLeuAspValMetValAlaLeuGlnGluSerSerGlyLeu 147
DB 469 ATCGTGTCTACCAACCGCTGCTGATGTCGATGTCGAGGTCCTGCGGAGCTCTGCGGTC 528
QY 148 ProHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMet 167
DB 529 CCGACCAACATGATCTCGGTATGCGCTGATGCTGACCTCTGCTGCTGCTGCTGCTGCTG 228
QY 168 IleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHis 187
DB 589 GTCCCGACGCGCTGCTGCTCTCCCGGACGCTGCGAGCCACCGCTCATCGGAGGAGGAG 248
QY 188 GlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGlu 207
DB 649 GCGGACTGCATGCTCCGCTTGTCCGCTACATTACCGTGAACGACTACCGGAGGAGGAG 267
QY 207 ValAlaGlyGlyValGluLeuValArgLeuGlyGlnGlySerAlaTyrTyrAlaProGly 247
DB 709 TTCTACCAAGGACGCGGTAGTCACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 268
QY 248 AlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCys 267
DB 829 GCATCCGCTGTCGCCATGCGCAACATCTTCTTGAACGACGAGAAAGCGGCTCATCCGCTG 888
QY 268 SerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysVal 287
DB 889 AGTGCTGACTGCACCGAGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 948
QY 288 IleGlyGlyArgGlyValGlyLysIleGluLeuGluLeuThrAlaGlnGluArgGln 307
DB 949 ATTGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1008
QY 308 GluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaLeuAsp 327
DB 1009 CAGTTCAGAGAGTCCGTCGACGAGCTCATCGGCTCAACAGGCGGTTGCTGCTTCAG 1068
QY 328 Ala 328
DB 1069 GCG 1071

```

Db 769 GTGCTGGCGGAGATCGTCCGCTTCTCGCCAGCGGTCCGCTTACTACGCCCGCC 828
Qy 248 AlaserAlaileGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCys 267
Db 829 GCATCCGCTGTCGCCATCGCAACATCTCTTGAACGACGAAAGCGCGTCATCCCGTGC 888
Qy 268 SerCysTyrLeuGlnGlnTyrGlyValGlnAenHisTyrLeuGlyValProCysVal 287
Db 899 AGTGTGTACTCAACGAGAGTACCGCTTGAAGGACATGTTTCATGTGTCCTCCCGCGTC 948
Qy 288 IleGlyGlyArgGlyValGlyLysIleIleGluLeuGluLeuThrAlaGlnGluArgGln 307
Db 949 ATTGAGGCGCGGCATCGAGCGGTCTATCGAGCTCGAGCTGAACGAGGAGAGAG 1008
Qy 308 GluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaLeuAsp 327
Db 1009 CAGTTCAGAGATCGTCGACGACGTCTATCGCGCTCAACAAGCGCGGTGCTGCTTCAG 1068
Qy 328 Ala 328
Db 1069 GCG 1071

RESULT 4

PCT-US94-03796-1

; Sequence 1, Application PC/TUS9403796

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: GENE ENCODING THE LACTATE DEHYDROGENASE

; NUMBER OF SEQUENCES: 15

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US94/03796

; FILING DATE: 06-APR-1994

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/046,160

; FILING DATE: 12-APR-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Deconti, Giulio A.

; REGISTRATION NUMBER: 31,503

; REFERENCE/DOCKET NUMBER: DCI-066CPPC

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 227-7400

; TELEFAX: (617) 227-5941

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 951 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..945

PCT-US94-03796-1

Alignment Scores:

Pred. No.:	3,438-92	Length:	951
Score:	831.50	Matches:	167
Percent Similarity:	69.0%	Conservative:	51
Best Local Similarity:	52.8%	Mismatches:	95
Query Match:	49.3%	Indels:	3
DB:	6	Gaps:	2

US-09-390-846-2 (1-330) x PCT-US94-03796-1 (1-951)

Qy 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
Db 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28

Db 10 AAAGCAAAAATCGTTTGTAGTTGGCTCAGGTATGATTGGAGGAGTAATGGCTACCTTAATT 69
Qy 29 SerLeuArgGluLeuGlyAspValValIlePheAspValValProAenMetProMetGly 48
Db 70 GTTCAGAAAAAATTTAGGAGATAGTATTGTTGCATATTGTAAGAAACATGCCACATGGA 129
Qy 49 LysAlaMetAspIleSerHisAenSerSerValValAspThrGlyIleThrValTyrGly 68
Db 130 AAAGCTTTAGATACATCTCATTAATGTTATGGCATATTCAAATTCGCAAGTAAGTGGT 189
Qy 69 SerAenSerTyrGluCysLeuLysGlyAlaAspValValIleIleThrAlaGlyIleThr 88
Db 190 TCAACACACTTATCAGCATTTGGCTGGAGCAGATGTAGTAATAGTAACACAGCTGGATTACC 249
Qy 89 LysIleProGlyLysSerAspLysGluTyrSerArgMetAspLeuLeuProValAsnIle 108
Db 250 AAGGCCCCAGGAAGAGTGAACAAGATGGAATAGAGATGATTATTATTCACATTAAACAAC 309
Qy 109 LysIleMetArgGluValGlyAlaIleLysSerTyrCysProAenAlaPheValIle 128
Db 310 AAGATTATGATTGAATTTGGTGTGCATATTAAAGAAATTTGCCAATGCTTTTATTATT 369
Qy 129 AsnIleThrAenProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeuPro 148
Db 370 GTTGTAAACAAACCCAGTAGATGTTATGGTACAAATTATTACATCAACATTCAGGTGTTCT 429
Qy 149 HisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMetIle 168
Db 430 AAAAACAAGATTATTGTTTGGTGTATTAGATACATCAAGATTGAAGATTATACATA 489
Qy 169 AlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHisGly 188
Db 490 TCTCAGAAATTAATGATGCCCAAGAGATGTAATGCCACACATTTGAGGTGTCATGGA 549
Qy 189 AspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGluPhe 208
Db 550 AATAAAATGTTCTTTTAAAAAGATACATTTACTGTAGGTGTTATCCCTTTACAAGAAATTT 609
Qy 209 ValLysLysGlyTrpIleLysGlnGluValAspIleValGlnLysThrLysVal 228
Db 610 ATTAATAACAAGTTAATTTCTGATGCTGAATTAAGAAGCTATATTGTAGAACTGTTAAT 669
Qy 229 AlaGlyGlyGluIleValArgLeuGlyGlnGlySerAlaTyrTyrAlaProGlyVala 248
Db 670 ACTGCATTAGAAATGTAACCTTA-----CATGCATCACCATATGTGACACAGCTGCT 723
Qy 249 SerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCysSer 268
Db 724 GCTATTATCGAAATGGCTGAATCTCTACTTAAAGATTGAAAAAAGATTATTAATTTGCTCA 783
Qy 269 CysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysValIle 288
Db 784 ACCTTGTTAGAAGGACAATATGGACACTCCGATATATTTCGGTGGTACACCTGTTGTTTA 843
Qy 289 GlyGlyArgGlyValGluLysIleIleGluLeuLeuThrAlaGlnGluArgGlnGlu 308
Db 844 GGTCTAATGTTGTTGAACAGTTATCGAATTACATTAATAGTAGAGAAAAAGGATAA 903
Qy 309 LeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAla 324
Db 904 TTTGATGAAGCCATAGCTGAACACTAAGAGAATG--AAGGCATTAGCT 948

RESULT 5

US-08-270-013B-1

; Sequence 1, Application US/08270013B

; Patent No. 5686294

; GENERAL INFORMATION:

; APPLICANT: Sogabe et al.

; TITLE OF INVENTION: PROTEIN HAVING HEAT-RESISTANT MALATE

; TITLE OF INVENTION: DEHYDROGENASE ACTIVITY

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leydig, Voit & Mayer, Ltd.

127 ValIleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGlnUserSergly 146
 1040 ATCCATCGTCTTGACGAACCGGTCGATGCATGACGTATACGGTCTTTAAGGAATCCGGA 1099
 147 LeuProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgArg 166
 1100 TTCCCGAANAACCGCTCATCGGCCAGTGGGGCTCTTGATATACGGCGCGCTTCCGACG 1159
 167 MetIleAlaAspIysLeuGluValSerProArgAspValGlnGlyMetValIleGlyVal 186
 1160 TTCGTGCCGAGAGCTGAACATTTCCGTAAGAAGATGTCACTGGGTGTGTTTAGCGCGC 1219
 187 HisGlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSer 206
 1220 CATGGCGATGACATGTCGCGCTCGTCCGCTACTCGTACGCCGCGGCATTCGCTCGAA 1279
 207 GluPheValIysGlyTrpIleIysGlnGluValAspAspIleValGlnIysThr 226
 1280 AAACACTCATTCGAAA-----GATCGTTTGGAGCCCATCGTTGAGCGGACG 1324
 227 LysValAlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaPro 246
 1325 CGCAAGGGCGCGGTGAATCTGCACTGTCTCGGCAACGGCAGCGCTACTACGACCG 1384
 247 GlyAlaSerAlaIleGlnMetAlaGluSerTyrLeuIysAspArgIysArgValMetVal 266
 1385 GCGCGCTCGCTTCTCGAAATGGTCGAAGCGATTTTCGAAAGACCAGCGCCGCAATTTGCGC 1444
 267 CysSerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrIleuGlyValProCys 286
 1445 GCGATCGCTTACCTTGAAGCGCAATACGGCTATGAAGGCATTTATTTGGCGGTGCGGACG 1504
 287 ValIleGlyClyArgGlyValGluIysIleIleGluLeuGluLeuThrAlaGlnGluArg 306
 1505 ATCCTCGGCGCAACGGCATCGAAGAAGTATCGAGCTCGAGCTGACCGAAGAGGAAAA 1564
 307 GlnGluLeuGlnGlySerIleAspGluValIysGluMetGlnIysAlaIle 323
 1565 GC-GCGCTCGCCAAATCCGTCGAATCCGTTAAAAATGTCATGCGCATCGTG 1614

RESULT 6
 US-08-838-418-1
 ; Sequence 1, Application US/08838418
 ; Patent No. 5744342
 ; GENERAL INFORMATION:
 ; APPLICANT: Socabe et al.
 ; TITLE OF INVENTION: PROTEIN HAVING HEAT-RESISTANT MALATE
 ; TITLE OF INVENTION: DEHYDROGENASE ACTIVITY
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
 ; STREET: Two Prudential Plaza, Suite 4900
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60601-6780
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/838,418
 ; FILING DATE: 17-MAR-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/270,013
 ; FILING DATE: 01-JUL-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 164701/1993
 ; FILING DATE: 02-JUL-1993
 ; ATTORNEY/AGENT INFORMATION:

NAME: Hoover, Allen E.
 REGISTRATION NUMBER: 37354
 REFERENCE/DOCKET NUMBER: 78339
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 616-5600
 TELEFAX: (312) 616-5700
 TELEX: (25)3533
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1912 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Bacillus stearothermophilus
 STRAIN: ATCC12016
 US-08-838-418-1

Alignment Scores:
 Pred. No.: 2,81e-76 Length: 1912
 Score: 706.00 Matches: 140
 Percent Similarity: 65.9% Conservative: 69
 Best Local Similarity: 44.2% Mismatches: 96
 Query Match: 41.8% Indels: 13
 DB: 3 Gaps: 3

US-09-390-846-2 (1-330) x US-08-838-418-1 (1-1912)

Qy 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
 Db 695 CGAAGAAATCTCGGTGATCGCGCGGATTACCGGGGGGCGACGCGGTTCCTTTTG 754
 Qy 29 SerLeuArgGluLeuGlyAspValValLeuPheAspVal-----ValProMetPro 46
 Db 755 GCGCAAAAGAGCTCGCGACGCTGCTGTGTCGATATTCCGCAGCTTGAGAACCAACG 814
 Qy 47 MetGlyValAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrVal 66
 Db 815 AAGGGAAGCGCTCGATATGCTCGAGGCAAGCCCGTCTCGGTTTGACCGCAATATC 874
 Qy 67 TyrGlySerAsnSerTyrGluCysLeuLysGlyAlaAspValValIleThrAlaGly 86
 Db 875 ATCGGCACATCGATTACGCTGACACAGCGGATTCGACATCGTCTCATCACAGCGC 934
 Qy 87 IleThrIleThrProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProVal 106
 Db 935 ATCGCCCGCAAGCGCGC-----ATGAGCGCGCGAGATTGCTGTGACGACG 979
 Qy 107 AsnIleLysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAlaPhe 126
 Db 980 AACCAAAATTTATGAAGCAAGTGACGAAAGGAGTCTGCTCAATATCTCGCGCAATCTAC 1039
 Qy 127 ValIleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerGly 146
 Db 1040 ATCATCTCTTGACGAACCCGCTGATGCGATGACGTATACGCTCTTTAAGCAATCCGA 1099
 Qy 147 LeuProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArg 166
 Db 1100 TTCCGAAAGAACCGCATCGCGCGAGTCCGGGCTCTTGGATAGCGCGCTTCGCGACG 1159
 Qy 167 MetIleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyVal 186
 Db 1160 TTCTGCCCGGAGGAGCTGAACATTTTCGGTAAAGATGTCTCACTGGGTTTGTAGCGCGC 1219
 Qy 187 HisGlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSer 206
 Db 1220 CATGGCGATGACATGGTGGCGCTCGTCCGCTACTCTGACGCGCGCGGCGCATTCGCTCGA 1279
 Qy 207 GluPheValLysLysGlyTyrIleLysGlnGluValAlaAspIleValGlnLysThr 226

NAME: Hoover, Allen E.
 REGISTRATION NUMBER: 37354
 REFERENCE/DOCKET NUMBER: 78339
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 616-5600
 TELEFAX: (312) 616-5700
 TELEX: (25)3533
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1912 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Bacillus stearothermophilus
 STRAIN: ATCC12016
 US-08-838-418-1

Alignment Scores:
 Pred. No.: 2,81e-76 Length: 1912
 Score: 706.00 Matches: 140
 Percent Similarity: 65.9% Conservative: 69
 Best Local Similarity: 44.2% Mismatches: 96
 Query Match: 41.8% Indels: 13
 DB: 3 Gaps: 3

US-09-390-846-2 (1-330) x US-08-838-418-1 (1-1912)

Qy 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
 Db 695 CGAAGAAATCTCGGTGATCGCGCGGATTACCGGGGGGCGACGCGGTTCCTTTTG 754
 Qy 29 SerLeuArgGluLeuGlyAspValValLeuPheAspVal-----ValProMetPro 46
 Db 755 GCGCAAAAGAGCTCGCGACGCTGCTGTGTCGATATTCCGCAGCTTGAGAACCAACG 814
 Qy 47 MetGlyValAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrVal 66
 Db 815 AAGGGAAGCGCTCGATATGCTCGAGGCAAGCCCGTCTCGGTTTGACCGCAATATC 874
 Qy 67 TyrGlySerAsnSerTyrGluCysLeuLysGlyAlaAspValValIleThrAlaGly 86
 Db 875 ATCGGCACATCGATTACGCTGACACAGCGGATTCGACATCGTCTCATCACAGCGC 934
 Qy 87 IleThrIleThrProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProVal 106
 Db 935 ATCGCCCGCAAGCGCGC-----ATGAGCGCGCGAGATTGCTGTGACGACG 979
 Qy 107 AsnIleLysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAlaPhe 126
 Db 980 AACCAAAATTTATGAAGCAAGTGACGAAAGGAGTCTGCTCAATATCTCGCGCAATCTAC 1039
 Qy 127 ValIleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerGly 146
 Db 1040 ATCATCTCTTGACGAACCCGCTGATGCGATGACGTATACGCTCTTTAAGCAATCCGA 1099
 Qy 147 LeuProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArg 166
 Db 1100 TTCCGAAAGAACCGCATCGCGCGAGTCCGGGCTCTTGGATAGCGCGCTTCGCGACG 1159
 Qy 167 MetIleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyVal 186
 Db 1160 TTCTGCCCGGAGGAGCTGAACATTTTCGGTAAAGATGTCTCACTGGGTTTGTAGCGCGC 1219
 Qy 187 HisGlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSer 206
 Db 1220 CATGGCGATGACATGGTGGCGCTCGTCCGCTACTCTGACGCGCGCGGCGCATTCGCTCGA 1279
 Qy 207 GluPheValLysLysGlyTyrIleLysGlnGluValAlaAspIleValGlnLysThr 226


```

; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2696
; LENGTH: 978
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2696

Alignment Scores:
Pred. No.: 9 37e-69 Length: 978
Score: 641.00 Matches: 132
Percent Similarity: 64.0% Conservative: 71
Best Local Similarity: 41.6% Mismatches: 102
Query Match: 38.0% Indels: 12
DB: 3 Gaps: 4

US-09-390-846-2 (1-330) x US-09-134-001C-2696 (1-978)
Qy 9 ArgProlylIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
Db 46 AGAAGAAAGATATCAATATTATGGCGCGGCACATACAGGTGGGACTCTAGCATTCATTCTT 105
Qy 29 SerLeuArgGluLeuGlyAspValValLeuPheAsp-----ValValProAsnMetPro 46
Db 106 GCACAAAAGGAATTAGGAGATATTGTGTGATTGAACGCCAGCAATCAGAGGGTATGGCT 165
Qy 47 MetGlyLysAlaMetAspIleSerHisAsnSerValValAspThrGlyIleThrVal 66
Db 166 AAAGAAAGGGTTAGATATTATTAGAAAGCGGACCATTTGGGGTTTGACACATCTGTA 225
Qy 67 TyrGlySerAsnSerTyrGluCysLeuLysGlyAlaAspValValIleIleThrAlaGly 86
Db 256 CATGTTTCAGTAAATATAGAGATATTAAAGATTACAGACATAGTGGTGCATGACGCAGGT 285
Qy 87 IleThrLysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuProVal 106
Db 286 -----ATACCTPAGGAATCAGGA-----ATGCAAGAGAAAGAAATTAGTTCAAACT 330
Qy 107 AsnIleLysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAlaPhe 126
Db 331 AATGAACAAATAGTACGAGAACTGCATTACAAATGCAACGTATGCACCTCATTCATA 390
Qy 127 ValIleAsnIleThrAsnProLeuAspValMetValAlaLeuGlnGluSerSerGly 146
Db 391 ATTATTGTATTGACTAATCCGGTTGATGTATGACATATACATATCTGCAATTTAAAGCATCAGGT 450
Qy 147 LeuProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArg 166
Db 451 TTTCTTAAGAACGTTATTATTGGTCAATCTGGAATTTTAGAGCTGCGAAGATATCGAAT 510
Qy 167 MetIleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyVal 186
Db 511 TTTATTGCTCAGAACTTAACGTGCTCTGCAAGATGTAATGGGTTGTTTGTAGTGA 570
Qy 187 HisGlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSer 206
Db 571 CATGGTGATACGATGTACCTTTGATTGATTAATACACACATTAATGCGGATCCAGTT--- 627
Qy 207 GluPheValLysGlyTrpIleLysGlnGluValAspAspIleValGlnLysThr 226
Db 628 -----AAGCATCTTATTTCTGAAGAAAGATTGATCAAAATTTGTTGAACGTACA 675
Qy 227 LysValAlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaPro 246
Db 676 CTTAAGGGTGGCGAAGAAATTTGCTTACTAGTCAAGGCTCAGCATATTATTCACCA 735
Qy 247 GlyAlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetVal 266
Db 736 GCAACTGCTATATGAACTATAGATGCAATTTTAAATGATCGGAAACGGTATTATACCA 795
Qy 267 CysSerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCys 286
Db 267 CysSerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCys 286

```

```

Db 796 AGTATTGCTTATCTAGAGGAGAAATACGGTTGTTTCAGATATTGTTTCGGAGTCTCTACT 855
Qy 287 ValIleGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGluArg 306
Db 856 ATAATAGGATATCAAGGAATAGAAAAGATTATAGAGGTAGATATGATATGATATGATGATAT 915
Qy 307 GlnGluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIle 323
Db 916 CAACAACACTACACACTCTGCGCAAGCTGTGAGTGAAGTCAAAACACTCACTA 966

RESULT 10
US-08-920-812-7
; Sequence 7, Application US/08920812
; Patent No. 5763188
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotaugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,812
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5024 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus epidermidis
; STRAIN: Clinical Isolate SE-32
US-08-920-812-7

Alignment Scores:
Pred. No.: 1.65e-66 Length: 5024
Score: 632.00 Matches: 132
Percent Similarity: 64.2% Conservative: 72
Best Local Similarity: 41.5% Mismatches: 101
Query Match: 37.4% Indels: 13
DB: 2 Gaps: 4

US-09-390-846-2 (1-330) x US-08-920-812-7 (1-5024)
Qy 9 ArgProlylIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
Db 57 AGAAGAAAGATATCAATATTATGGCGCGGCACATACAGGTGGGACTCTAGCATTCATTCTT 116

```

QY 29 SerLeuArgGluLeuGlyAspValValLeuPheAsp-----ValValProAsnMetPro 46
Db 117 GCACAAAGGAATAGGAGATATTGTGTGATTGAACGCCAGCAATCAGAGGGTATGGCT 176
QY 47 MetGlyLysAlaMetAspHisAsnSerValValAspThrGlyLeuThrVal 66
Db 177 AAGGAAGGGCTAGATATTATTAAGAACGGGACCCATTGGGGTTTGACACATCTGTA 236
QY 67 TyrGlySerAsnSerTyrGluCysLeuLysGlyAlaAspValValLeuThrAlaGly 86
Db 237 CATGTTTCAGTAAATATAGACATATTAAAGATTGAGACATAGTGGTATGATGACGCT 296
QY 87 IleThrLysIleProGlyLysSerAspLysGluTrpSerArgMet-AspLeuLeuProVa 106
Db 297 -----ATACCTAGGAATCAGGA-----ATGACAAGGAGGAAGAAATAGTTCAAAC 341
QY 106 LasnIleLysIleMetArgGluValGlyAlaLeuLysSerTyrCysProAsnAlaPh 126
Db 342 TAATGACAATAATAGCAGAAACTGCTATCAAAATGCAAGTATGACCTCATTCAT 401
QY 126 eValIleAsnIleThrAsnProLeuAspValMetValAlaLeuGlnGluSerSerG1 146
Db 402 AATTATTGTTAGCTAATCCGGTTGATGATGACATATATGATGATGATGATGATGATGAT 461
QY 146 YLeuProHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgAr 166
Db 462 TTTTCTTAAGAACGCTATTATTGTTCAATCTGGAATTTAGACGTCGCAAGATATGCAAC 521
QY 166 gMetIleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyVa 186
Db 522 TTTTATTCTCAAGAACCTAACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 581
QY 186 lHisGlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSe 206
Db 582 ACATGGTATACGATGTTACCTTTGATTAAATACACACACATTAATGGGATTCAGTT-- 639
QY 206 rGluPheValLysLysGlyTrpIleLysGlnGluValAspIleValGlnLysTh 226
Db 640 -----AAGCATCTATTCTGCAAGAAAGATGATCAAAATGTTGAACGTPAC 686
QY 226 rLysValAlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaPr 246
Db 687 ACCTAAGGGTGGTGCAGAAATTTGTCATTAATGCTAGGTCAGGTCAGATATTATGACC 746
QY 246 oGlyAlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetVa 266
Db 747 AGCACTGCTATATATGAACATATAGATGCTAATTTTAAATGATCGAAACGGTTATTACC 806
QY 266 lCysSerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCy 286
Db 807 AAGTATTGCTTATCTAGAGGAGAAATACGGTTGTTTTCAGATATTGTTTCGGAGTTCTTAC 866
QY 286 sValIleGlyArgGlyValGluLysIleLeuLeuGluLeuLeuThrAlaGlnLuar 306
Db 867 TATATAGGATATCAAGGAATAGAAAGATATATAGAGGTAGATAGATGAATAATGATGATGA 926
QY 306 gGlnGluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIle 323
Db 927 TCACAACTACAACTCTGCGCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 978

RESULT 11

US-08-920-827-7

; Sequence 7, Application US/08920827

; Patent No. 5770375

; GENERAL INFORMATION:

; APPLICANT: Ohno, Tsuneya

; APPLICANT: Matsuhisa, Akio

; APPLICANT: Uehara, Hirotsugu

; APPLICANT: Eda, Soji

; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease

; NUMBER OF SEQUENCES: 25

; CORRESPONDENCE ADDRESS:

;

;

;

;

;

;

;

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/920,827
FILING DATE: 29-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,577
FILING DATE: 27-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 5024 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORGANISM: Staphylococcus epidermidis
STRAIN: Clinical Isolate SE-32
US-08-920-827-7

Alignment Scores:

Pred. No.:	1.65e-66	Length:	5024
Score:	632.00	Matches:	132
Percent Similarity:	64.3%	Conservative:	72
Best Local Similarity:	41.5%	Mismatches:	101
Query Match:	37.4%	Indels:	13
DB:	2	Gaps:	4

US-09-390-846-2 (1-330) x US-08-920-827-7 (1-5024)

QY 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
Db 57 AGAAGAAAGATATCAATTTATTGGCGGGACATACAGGTGGGACTCTAGCATTCATTCTT 116
QY 29 SerLeuArgGluLeuGlyAspValValLeuPheAsp-----ValValProAsnMetPro 46
Db 117 GCACAAAGGAATTAGGAGATATTGTGTGATTGAACGCCAGCAATCAGAGGGTATGGCT 176
QY 47 MetGlyLysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyLeuThrVal 66
Db 177 AAGGAAGGGCTAGATATTATTAAGAACGGGACCCATTGGGGTTTGACACATCTGTA 236
QY 67 TyrGlySerAsnSerTyrGluCysLeuLysGlyAlaAspValValIleThrAlaGly 86
Db 237 CATGTTTCAGTAAATATAGACATATTAAAGATTGAGACATAGTGGTATGATGACGCT 296
QY 87 IleThrLysIleProGlyLysSerAspLysGluTrpSerArgMet-AspLeuLeuProVa 106
Db 297 -----ATACCTAGGAATCAGGA-----ATGACAAGGAGGAAGAAATAGTTCAAAC 341
QY 106 lAsnIleLysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAlaPh 126
Db 342 TAATGACAATAATAGCAGAAACTGCTATCAAAATGCAAGTATGACCTCATTCAT 401
QY 126 eValIleAsnIleThrAsnProLeuAspValMetValAlaLeuGlnGluSerSerG1 146

```

;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5024 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORGANISM: Staphylococcus epidermidis
; STRAIN: Clinical Isolate SE-32
; US-08-921-177-7

Alignment Scores: 1.65e-66 Length: 5024
Pred. No.: 632.00 Matches: 132
Score: 64.2% Conservative: 72
Percent Similarity: 41.5% Mismatches: 101
Best Local Similarity: 37.4% Indels: 13
Query Match: 2 Gaps: 4
DB:

US-09-390-846-2 (1-330) x US-08-921-177-7 (1-5024)

Qy 9 ArgProlysisIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
Db 57 AGAAGAAAGATATCAATTATTGGCGGGACATACAGGTGGGACTCTAGCATTCATTCTT 116
Qy 29 SerLeuArgGluLeuGlyAspValValPheAsp-----ValValProAsnMetPro 46
Db 117 GCACAAAGGAATTAGAGATATTTGTTGTAATGAACGCCAGCATCAGGGGTATGGCT 176
Qy 47 MetGlyLysAlaMetAspIleSerHisAsnSerValValAspThrGlyIleThrVal 66
Db 177 AAAGAAAGCGGTAGATATTTTAGAAAGCGGACCCATTTGGGGGTTTGACACATCTGTA 236
Qy 67 TyrGlySerAsnSerTyrGluCysLeuGlyAlaAspValValIleIleThrAlaGly 86
Db 237 CATGGTTTCAGTAATATAGAGATATTTAAAGATTCAGACATAGTGGTGATGACGTGCA 296
Qy 87 IleThrLysIleProGlyLysSerAspLysGluTyrSerArgMet-AspLeuLeuProVa 106
Db 297 -----ATACCTAGGAATCAGGA-----ATGACAGGAGAGAAGATTTGTTCAAAC 341
Qy 106 IAsnIleLysIleMetArgGluValGlyAlaIleLysSerTyrCysProAsnAlaPh 126
Db 342 TAATGAACAAATAGTACGAGAACTGCATTTACAAATTTGCAACGTTATGCACCTCATCAAT 401
Qy 126 eValIleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGl 146
Db 402 AATTATTGTATTGACTAATCCGTTGATGTATGACATATATCTGCATTTAAAGCATCAGG 461
Qy 146 YLeuProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgAr 166
Db 462 TTTTCTTAAAGAACGTTATTTTGTCTCAATCTGGAAATTTTAGACGCTCGAAGATATCGAAC 521
Qy 166 gMetIleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyVa 186
Db 522 TTTTATTGCTCAAGAACTTAACGTTCTGTCAAAGATGTAAATGGGTTTGTGGTGGTGG 581
Qy 186 lHisGlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSe 206
Db 582 ACATGGTGATACGATGTTTACCTTTTGATTAATAACACACACATTAATCGGATTCAGTT-- 639
Qy 206 rGluPheValLysLysGlyTyrIleLysGlnGluValAspAspIleValGlnLysTh 226
Db 640 -----AAGCATCTTATTTCTGAAGAAAGATGATCAAAATTTGTTGAACGTAC 686
Qy 226 rLysValAlaGlyGlyIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaPr 246
Db 687 ACGTAAGGGTGTGCAGAAATTTGTCATTACTAGTCAAGCTCAGCATATTATGACACC 746
Qy 246 oGlyAlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetVa 266
Db 747 AGCAACTCTATATATGAACACTATAGATGCAATTTTAAATGATCGGAACCGTTATTACC 806
Qy 266 lCysSerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCy 286
Db 807 AAGTATTGCTTATCTAGAGGAGAAATACGGTTGTTTTCAGATATTGTTTCCGAGTTCCTAC 866
Qy 286 sValIleGlyGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGluAr 306
Db 867 TATTAATAGGATATCAAGGAATAGAAAGATTTATAGAGTAGATATGAATAATGATGAGTA 926
Qy 306 gGlnGluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIle 323
Db 927 TCACAACTACACACTCTCGCGCAAGATGTGAGTGAAGTCAAAACTCACTA 978

RESULT 12
US-08-921-177-7
; Sequence 7, Application US/08921177
; Patent No. 5798211
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/921.177
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420

```

```

;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5024 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORGANISM: Staphylococcus epidermidis
; STRAIN: Clinical Isolate SE-32
; US-08-921-177-7

Alignment Scores: 1.65e-66 Length: 5024
Pred. No.: 632.00 Matches: 132
Score: 64.2% Conservative: 72
Percent Similarity: 41.5% Mismatches: 101
Best Local Similarity: 37.4% Indels: 13
Query Match: 2 Gaps: 4
DB:

US-09-390-846-2 (1-330) x US-08-921-177-7 (1-5024)

Qy 9 ArgProlysisIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
Db 57 AGAAGAAAGATATCAATTATTGGCGGGACATACAGGTGGGACTCTAGCATTCATTCTT 116
Qy 29 SerLeuArgGluLeuGlyAspValValPheAsp-----ValValProAsnMetPro 46
Db 117 GCACAAAGGAATTAGAGATATTTGTTGTAATGAACGCCAGCATCAGGGGTATGGCT 176
Qy 47 MetGlyLysAlaMetAspIleSerHisAsnSerValValAspThrGlyIleThrVal 66
Db 177 AAAGAAAGCGGTAGATATTTTAGAAAGCGGACCCATTTGGGGGTTTGACACATCTGTA 236
Qy 67 TyrGlySerAsnSerTyrGluCysLeuGlyAlaAspValValIleIleThrAlaGly 86
Db 237 CATGGTTTCAGTAATATAGAGATATTTAAAGATTCAGACATAGTGGTGATGACGTGCA 296
Qy 87 IleThrLysIleProGlyLysSerAspLysGluTyrSerArgMet-AspLeuLeuProVa 106
Db 297 -----ATACCTAGGAATCAGGA-----ATGACAGGAGAGAAGATTTGTTCAAAC 341
Qy 106 IAsnIleLysIleMetArgGluValGlyAlaIleLysSerTyrCysProAsnAlaPh 126
Db 342 TAATGAACAAATAGTACGAGAACTGCATTTACAAATTTGCAACGTTATGCACCTCATCAAT 401
Qy 126 eValIleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGl 146
Db 402 AATTATTGTATTGACTAATCCGTTGATGTATGACATATATCTGCATTTAAAGCATCAGG 461
Qy 146 YLeuProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgAr 166
Db 462 TTTTCTTAAAGAACGTTATTTTGTCTCAATCTGGAAATTTTAGACGCTCGAAGATATCGAAC 521
Qy 166 gMetIleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyVa 186
Db 522 TTTTATTGCTCAAGAACTTAACGTTCTGTCAAAGATGTAAATGGGTTTGTGGTGGTGG 581
Qy 186 lHisGlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSe 206
Db 582 ACATGGTGATACGATGTTTACCTTTTGATTAATAACACACACATTAATCGGATTCAGTT-- 639
Qy 206 rGluPheValLysLysGlyTyrIleLysGlnGluValAspAspIleValGlnLysTh 226
Db 640 -----AAGCATCTTATTTCTGAAGAAAGATGATCAAAATTTGTTGAACGTAC 686
Qy 226 rLysValAlaGlyGlyIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaPr 246
Db 687 ACGTAAGGGTGTGCAGAAATTTGTCATTACTAGTCAAGCTCAGCATATTATGACACC 746
Qy 246 oGlyAlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetVa 266
Db 747 AGCAACTCTATATATGAACACTATAGATGCAATTTTAAATGATCGGAACCGTTATTACC 806
Qy 266 lCysSerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCy 286
Db 807 AAGTATTGCTTATCTAGAGGAGAAATACGGTTGTTTTCAGATATTGTTTCCGAGTTCCTAC 866
Qy 286 sValIleGlyGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGluAr 306
Db 867 TATTAATAGGATATCAAGGAATAGAAAGATTTATAGAGTAGATATGAATAATGATGAGTA 926
Qy 306 gGlnGluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIle 323
Db 927 TCACAACTACACACTCTCGCGCAAGATGTGAGTGAAGTCAAAACTCACTA 978

RESULT 12
US-08-921-177-7
; Sequence 7, Application US/08921177
; Patent No. 5798211
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/921.177
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420

```

QY 246 oGlyAlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetVa 266
 Db 747 AGCACTGCTATATATGAACATATAGATCAATTTTAAATGATCGGAACGGTTATTACC 806
 QY 266 lCysSerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCy 286
 Db 807 AGATTGCTTATCTAGAGGAGATACGGTTGTTTCAGATATTTGTTTCGGAGTTCCTAC 866
 QY 286 sValleGlyGlyArgGlyValGluLysIleleGluLeuGluLeuThrAlaGlnGluAr 306
 Db 867 TATAATAGGATATCAAGGAATAGAAAGATATAGAGGTAGATGAATAATGATGAGTA 926
 QY 306 gGlnGluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIle 323
 Db 927 TCACAACTACACACTCTCGCGAAGATGTGAGTGAAGTCAAAAACCTCACTA 978

RESULT 13

US-08-362-577C-7
 ; Sequence 7, Application US/08362577C
 ; Patent No. 5807673
 ; GENERAL INFORMATION:
 ; APPLICANT: Ohno, Tsuneya
 ; APPLICANT: Matsuhisa, Akio
 ; APPLICANT: Uehara, Hirotsugu
 ; APPLICANT: Eda, Soji
 ; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/362,577C
 ; FILING DATE: 27-MAR-1995
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Rin-Laures, Li-Hsien
 ; REGISTRATION NUMBER: 33,547
 ; REFERENCE/DOCKET NUMBER: 19036/32420
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/474-6300
 ; TELEFAX: 312/474-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 5024 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: Genomic DNA
 ; ORIGINAL SOURCE:
 ; ORGANISM: Staphylococcus epidermidis
 ; STRAIN: Clinical Isolate SE-32
 ; US-08-362-577C-7

Alignment Scores:
 Pred. No.: 1.65e-66 Length: 5024
 Score: 632.00 Matches: 132
 Percent Similarity: 64.2% Conservative: 72
 Best Local Similarity: 41.5% Mismatches: 101
 Query Match: 37.4% Indels: 13
 DB: 2 Gaps: 4

US-09-390-846-2 (1-330) x US-08-362-577C-7 (1-5024)

QY 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyGlyThrMetAlaPheLeuCys 28
 Db 57 AGAAGAAAGATATCAATATTATGGCGCGGACATACAGGTGGGACTCTAGCATTCATTCTT 116
 QY 29 SerLeuArgGluLeuGlyAspValValLeuPheAsp-----ValValProAsnMetPro 46
 Db 117 GCACAAAGGAATTAGGAGATATTGTGTGATTGAACCGCAGCAATCAGAGGGTATGGGT 176
 QY 47 MetGlyLysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrVal 66
 Db 177 AAAGAAAGGCGTTAGATATTTTAGAAAGCGGACCCATTTCGGGGTTCGACACATCTGTA 236
 QY 67 TyrGlySerAsnSerTyrGluCysLeuLysGlyAlaAspValValIleleThrAlaGly 86
 Db 237 CATGTTTCAGTAAATATAGAAATATTAAGATTTCAGACATAGTGTGATGATCAGGT 296
 QY 87 IleThrLysIleProGlyLysSerAspLysGluTrpSerArgMet-AspLeuLeuProVa 106
 Db 297 -----ATACCTAGGAATCAGGA-----ATGACAAGGAGGAAGAAATTAGTTCAAAC 341
 QY 106 LasnIleLysIleMetArgGluValGlyAlaIleLysSerTyrCysProAsnAlaPh 126
 Db 342 TAATCAACAATAGTACGAGAACTGCAATTACAATTCACAGTATGCACCTCATTCAT 401
 QY 126 eValIleAsnIleThrAsnProLeuAspValMetValAlaLeuGlnGluSerSerGl 146
 Db 402 AATTATTGTTACTAATCCGGTTGATGTATGACATATATCTGATGATTTAAAGCATCAGG 461
 QY 146 YLeuProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgAr 166
 Db 462 TTTTCTAAAGAACGTTATTATTGGTCAATCTGGAATTTTAGACGCTCGAAGATATCGAAC 521
 QY 166 gMetIleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyVa 186
 Db 522 TTTTATTGCTCAAGAACTTAACGTGCTCTCAAGATGTAATGGTTGTTTATAGTGG 581
 QY 186 lHisGlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSe 206
 Db 582 ACATGGTGATCAGATGTTTACCTTTGATTATAACACACACACATTAATGGGATTCAGATT-- 639
 QY 206 rGluPheValLysLysGlyTrpIleLysGlnGluValAspAspIleValGlnLysTh 226
 Db 640 -----AAGCATCTTATTCTGAAGAAAAGATTGATCAAAATGTTGTAACGTAC 686
 QY 226 rLysValAlaGlyGlyGluIleValArgLeuGlyGlnGlySerAlaTyrTyrAlaPr 246
 Db 687 AGTAAGGTGTCAGAAATTTGTCATTACTAGTCAAGCTCAGCATATTATGCAACC 746
 QY 246 oGlyAlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetVa 266
 Db 747 AGCAACTGCTATATATGAACATATAGATCAATTTTAAATGATCGGAACGGTTATTACC 806
 QY 266 lCysSerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCy 286
 Db 807 AAGTATTGCTTATCTAGAGGAGAAATGTTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTT 866
 QY 286 sValIleGlyArgGlyValGluLysIleleGluLeuGluLeuThrAlaGlnGluAr 306
 Db 867 TATAATAGGATATCAAGGAATAGAAAGATATAGAGGTAGATGAATAATGATGAGTA 926
 QY 306 gGlnGluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIle 323
 Db 927 TCACAACTACACACTCTCGCGAAGATGTGAGTGAAGTCAAAAACCTCACTA 978

RESULT 14

US-08-920-828-7
 ; Sequence 7, Application US/08920828
 ; Patent No. 5853998
 ; GENERAL INFORMATION:
 ; APPLICANT: Ohno, Tsuneya
 ; APPLICANT: Matsuhisa, Akio
 ; APPLICANT: Uehara, Hirotsugu

```

; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,828
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5024 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus epidermidis
; STRAIN: Clinical Isolate SE-32
US-08-920-828-7

```

```

Alignment Scores:
Pred. No.: 1,658-66 Length: 5024
Score: 632.00 Matches: 132
Percent Similarity: 64.2% Conservative: 72
Best Local Similarity: 41.5% Mismatches: 101
Query Match: 37.4% Indels: 13
DB: 2 Gaps: 4

```

US-09-390-846-2 (1-330) x US-08-920-828-7 (1-5024)

```

Qy 9 ArgProLyAlaMetValGlySerGlyMetIleGlyGlyThrMetAlaPheLeuCys 28
Db 57 AGAAGAGATATCAATATTGCGCGGCACATACAGTGGGACTTAGCATTCATTCTT 116
Qy 29 SerLeuArgGluLeuGlyAspValValLeuPheAsp-----ValValProAsnMetPro 46
Db 117 GCACAAAGGAATATTAGGAGATATTGTTGATTGAAGCCAGCAATCAGAGGGTATGGCT 176
Qy 47 MetGlyLyAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrVal 66
Db 177 AAGAGAAAGCGCTTAGATATTATTTAGAAAGCGGACCCATTGGGGGTTTGACACATCTGTA 236
Qy 67 TyrGlySerAsnSerTyrGlyCysLeuLyGlyAlaAspValValIleThrAlaGly 86
Db 237 CATGGTTCAGTAATATAGAAAGATATTAAAGATTACAGATAGTGGTGTAGTACGAGT 296
Qy 87 IleThrLyAlaProGlyLySerAspLyGluTrpSerArgMet-AspLeuLeuProVa 106
Db 297 -----ATACCTAGGAAATCAGGA-----ATGACAAAGGAGAGAATATTAGTTCAAAC 341
Qy 106 lAsnIleLyAlaMetArgGluValGlyAlaIleLySerTyrCysProAsnAlaPh 126

```

```

Db 342 TAATGAACAATAGTACGAGAACTGCATTACAAATTTGCAACGTATGCACCTCATTCAT 401
Qy 126 eValIleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGl 146
Db 402 AATTATTGTATTGACTAATCCGGTTGATGTATGACATATATGTCATTTAAAGCATCAGG 461
Qy 146 YLeuProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgAr 166
Db 462 TTTTCTTAAAGAACGTATTATTGTCATCTGGAATTTTAGACCTCGAAGATATCGAAC 521
Qy 166 gMetIleAlaAspLyLeuGluValSerProArgAspValGlnGlyMetValIleGlyVa 186
Db 522 TTTTATTGCTCAAGAACTTAACGTGCTGTCACAAAGATGTAATGGTGTGTTTAGGTGG 581
Qy 186 lHisGlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSe 206
Db 582 ACATGGTGATACGATGTTACCTTTTGATTAATAACACACACACATTAATGGGATTCAGT 639
Qy 206 rGluPheValLySlyGlyTrpIleLySglnGluValAspAspIleValGlnLyvTh 226
Db 640 -----AAGCATCTTATTTCTGAAGAAAGATGATCAAAATTGTTGACGTAC 686
Qy 226 rLySValAlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaPr 246
Db 687 ACGTAAAGGGTGGTCAGAAATTTGTCATTACTAGGTCAAGGCTCAGCATATTATGCACC 746
Qy 246 oGlyAlaSerAlaIleGlnMetAlaGluSerTyrLeuLySAspArgLySArgValMetVa 266
Db 747 AGCAACTGCTATATATGAAACTATAGATGCAATTTTAAATGATCGAAACGGTTATTACC 806
Qy 266 lCysSerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCy 286
Db 807 AAGTATTGCTTATCTAGGAGGAATACGGTTGTTTCAGATATTGTTTCGAGATTCCTAC 866
Qy 286 sValIleGlyGlyArgGlyValGluLySlyIleIleGluLeuGluLeuThrAlaGlnGluAr 306
Db 867 TATAATAGGATATCAAGGAATAGAAAGATTATAGAGGTAGATATGAATATGATGAGTA 926
Qy 306 gGlnGluLeuGlnGlySerIleAspGluValLySglnMetGlnLySAlaIle 323
Db 927 TCAACAACACTACACACTCTGCGCAAGATGTGAGTGAAGTCAAAACTCACTA 978

```

RESULT 15

```

US-09-107-532A-2832
; Sequence 2832, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310

```

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

; NAME: Ariniello, Pamela Deneke
 ; REGISTRATION NUMBER: 40,489
 ; REFERENCE/DOCKET NUMBER: GTC-012
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (781)893-5007
 ; TELEFAX: (781)893-8277
 ; INFORMATION FOR SEQ ID NO: 2832:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 975 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: circular
 ; MOLECULE TYPE: DNA (genomic)
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Enterococcus faecium
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (8) LOCATION 1...975
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2832:
 US-09-107-532A-2832

Alignment Scores:
 Pred. No.: 5,348-47 Length: 975
 Score: 464.00 Matches: 106
 Percent Similarity: 55.1% Conservative: 72
 Best Local Similarity: 32.8% Mismatches: 135
 Query Match: 27.5% Indels: 10
 DB: 3 Gaps: 5

US-09-390-846-2 (1-330) x US-09-107-532A-2832 (1-975)

QY 3 ValPheGluLysAsnThrArgProLysIleAlaMetValGlySerGlyMetIleGly 22
 Db 25 TTATTCATGAAAACAAAGTCGTAAGTAGTAGTAATCGTCGGTACAGGTTTGTGGACA 84
 QY 23 ThrMetAlaPheLeuCysSerLeuArgGluLeu---GlyAspValValLeuPheAspVal 41
 Db 85 AGTATCGCTTATGCAATCATCAACCAAGGAATCTCTAATGAATTAGTATTGATCGATGTG 144
 QY 42 ValProAsnMetProMetGlyLysAlaMetAspIleSerHisAsnSerSerValValAsp 61
 Db 145 AATCAAGAAAACGAGAGGGAAGCATTTGGACCTTTTAGATGCTTGGGGCGCAT 204
 QY 62 ThrGlyIleThrValTyrGlySerAsnSerTyrGluCysLeuLysGlyAlaAspValVal 81
 Db 205 GAAATGTGGCTGTATGG---TCAGGTGGCTATGAGAAATGTAAAGATGCTGATATCGTT 261
 QY 82 IleIleThrAlaGlyIleThrLysIleProGlyLysSerAspLysGluTrpSerArgMet 101
 Db 262 GTTATCACAGCTGGTATCAACCAAAACCTGGTCAA-----TCTCGTTTA 306
 QY 102 AspLeuLeuProValAsnIleLysIleMetArgGluValGlyAlaAlaIleLysSerTyr 121
 Db 307 GATTTAGTTAAACAAATCGTCATATCGGCCAAATCGTAAAGAAATCATGGGATCA 366
 QY 122 CysProAsnAlaPheValIleAsnIleThrAsnProLeuAspValMetValAlaAlaLeu 141
 Db 367 GGATTCGACGGTATTATGTTGCTTCAACCCAGTAGATATTTTGACATATATCGCT 426
 QY 142 GlnGluSerSerGlyLeuProHisArgIleCysGlyMetAlaGlyMetLeuAspSer 161
 Db 427 TGGACGAATCTGGCTGCCACATCATCGTGTATCGGTACTGGAAACAACATTGGATACA 486
 QY 162 SerArgPheArgMetIleAlaAspLysLeuGluValSerProArgAspValGlnGly 181
 Db 487 ACTCGTTTCCGTAAGAAATCGCATTGAAATTAAGTTGACCCACGAGCGGTCCACGGC 546
 QY 182 MetValIleGlyValHisGlyAspHisMetValProLeuSerArgTyrAlaThrValAsn 201
 Db 547 TATATCTTAGGTGAACATGGTGATTCAGAAAGTTGTCGATGTCATACATCAACTGTCGGT 606

QY 202 GlyIleProLeuSerGluPheValLysLysGlyTrp---IleLysGlnGluValValAsp 220
 Db 607 GGTAAACCAAGTGTTTGAAATTTAGAAAAAGACCATCGTATTGCCCCAAGATGAATTAGAT 666
 QY 221 AspIleValGlnLysThrLysValAlaGlyGlyGluIleValArgLeuLeuGlyGlnGly 240
 Db 667 GTGATTGCTGATAAAGTCGGTAATGCTCTTATGAGATCATTTGAT-----CGTAAAAA 720
 QY 241 SerAlaTyrTyrAlaProGlyAlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAsp 260
 Db 721 GCAACTTACTATGTTGTTGTTGATGAGTACTGCACGTCATCGTAAAGCTATCTTGAATAAC 780
 QY 261 ArgLysArgValMetValCysSerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHis 280
 Db 781 GAACAAGCTGTATTGCTCTATCTCTATTAACTGGTGAATACGACGAAAAAGATATC 840
 QY 281 TyrLeuGlyValProCysValIleGlyArgGlyValGluLysIleIleGluLeuGlu 300
 Db 841 TTTACAGGTGTCCTCATCAATCGTTGACGAAATGGTGTTCGTGAAGTTGTTGAATTATCA 900
 QY 301 LeuThrAlaGlnGluArgGlnGlnLeuGlnGlySerIleAspGluValLysGluMetGln 320
 Db 901 ATCAACGAAGAAGAAAAAGCAATGTTCAACAAATCAACAAGTGTCTTTAAGAGAGATATTG 960
 QY 321 LysAlaIle 323
 Db 961 AACACTGTA 969

Search completed: March 3, 2006, 01:19:18
 Job time : 221 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 2, 2006, 23:54:34 ; Search time 3708 Seconds
(without alignments)
4163.899 Million cell updates/sec

Title: US-09-390-846-2

Perfect score: 1688

Sequence: 1 MAVFKNTRPKIANVGSGLI.....GSIDEVKEMQKATAALDASK 330

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/abs/ABSSWEB.spool/US09390846/runat_02032006_104232_4868/app.query.fasta_1
-DB=EST -OPMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs06p
-USER=US09390846 @CGN 1.1 8010 @runat_02032006_104232_4868 -ICPU=3
-NO MMAP -NEG SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOF=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*
10: gb_gses.*
11: gb_gses3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	769	45.6	811	7	CV643661
2	762	45.1	804	7	CV637221
3	742.5	44.0	784	6	CB025968
4	741	43.9	777	7	CV637428
5	721.5	42.7	753	6	CB025472
6	718	42.5	900	8	CB022234
7	702	41.6	684	7	CV551869

8	686.5	40.7	772	6	CB022051
9	683.5	40.5	780	6	CB030949
10	674.5	40.0	721	6	CB027630
11	671.5	39.8	630	6	CF259970
12	659	39.0	586	6	CF268212
13	658	39.0	628	7	CN621209
14	656.5	38.9	744	6	CB025686
15	656.5	38.9	746	7	CK734706
16	653	38.7	585	6	CB0301250
17	651.5	38.6	766	6	CB024558
18	650	38.5	611	6	CB754431
19	645	38.2	595	6	CF659673
20	641.5	38.0	752	6	CB023734
21	636.5	37.7	577	6	CF268385
22	623	36.9	602	6	CF421715
23	621.5	36.8	714	7	CK734359
24	620	36.7	691	6	CB065846
25	619.5	36.7	580	6	CF341663
26	618	36.6	746	8	CO221116
27	613.5	36.3	582	6	CB411873
28	613.5	36.3	615	7	CN122464
29	611.5	36.2	576	6	CB301358
30	611	36.2	690	3	BM169276
31	608	36.0	577	7	CF967655
32	607	36.0	571	7	CO053115
33	605.5	35.9	709	7	CF702488
34	601.5	35.6	580	6	CF268206
35	601.5	35.6	582	7	CF942651
36	598	35.4	770	1	AJ821572
37	590	35.0	559	6	CB367907
38	588.5	34.9	561	3	BM175091
39	587	34.8	556	6	CF247444
40	573	33.9	817	7	CV647412
41	565.5	33.5	537	3	BM271642
42	562.5	33.3	716	7	CV549539
43	562	33.3	535	7	CN196747
44	560.5	33.2	526	6	CB187397
45	558	33.1	576	7	CN621319

ALIGNMENTS

RESULT 1
CV643661
LOCUS
DEFINITION
EST922414 Field isolate cDNA library Plasmodium vivax cDNA clone
PVMZ82 5' end, mRNA sequence.
CV643661
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CV643661 811 bp mRNA linear EST 01-JAN-2005
EST922414 Field isolate cDNA library Plasmodium vivax cDNA clone
PVMZ82 5' end, mRNA sequence.
CV643661 GI:56950479
EST.
Plasmodium vivax (malaria parasite P. vivax)
Plasmodium vivax
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
1 (bases 1 to 811)
Carlton,J.M. and Cui,L.
A survey of genes in Plasmodium vivax by EST sequencing
Unpublished (2004)
Contact: Jane Carlton
Parasite Genomics Group
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-530-9319
Fax: 301-838-0208
Email: carlton@tigr.org
Seq primer: T1.
Location/Qualifiers
1. .811
/organism="Plasmodium vivax"
/mol_type="mRNA"
/strain="Field isolate"
/db_xref="taxon:5855"
/clone_lib="Field isolate cDNA library"

FEATURES
source


```

Db      305  GGACATATTAGAACCTTTGGCCCAATGCCCTTTATCATTTGGTGACAAACCCAGTGGAC 364
Qy      136  ValMetValAlaLaLeuGlnGluSerSerGlyLeuProHisHisArgIleCysGlyMet 155
Db      365  GTGATGTGCGAGTTACTCTTCGAGCATTTCCGGAGTCCCAAAAATAAATCATCGGATTA 424
Qy      156  AlaGlyMetLeuAapSerSerArgPheArgMetIleAlaAapLysLeuGluValSer 175
Db      425  GGTGGTGTGTAGATACATCTAGACTGAATATTTACATATCGCAGAAAGTTGAACGTCCGC 484
Qy      176  ProArgAapValGlnGlyMetValIleGlyValHisGlyAapHisMetValProLeuSer 195
Db      485  CCGAGAGATGTTAATGCACCTATTCGTGGTCCATCGGACAGATGGTTCCTCTGAAA 544
Qy      196  ArgTyrAlaThrValAanGlyIleProLeuSerGluPheValLysLysGlyTyrIleLys 215
Db      545  AGGTACATCAGTTGAGGTATCCCATTCGCAAGATTTATTATAACAAAAGATTACA 604
Qy      216  GlnGluGluValAapAapIleValGlnLysThrLysValAlaGlyGlyGluIleValArg 235
Db      605  GATGAAGAAGTGAAGCATATTTGATCGCACTGTGAACACTGCTTTTCGAGATTGTGAGC 664
Qy      236  LeuLeuGlyGlnGlySerAlaTyrTyrAlaProGlyAlaSerAlaIleGlnMetAlaGlu 255
Db      665  CTCCTT-----GCCTCTCCTTATGTTGCCCACTGCTGCCATCATCGAAATGCCGAA 718
Qy      256  SerTyrLeuLysAapArgLysArgValMetValCysSerCysTyrLeuGlnGlyGlnTyr 275
Db      719  TCTATTGGAAGGATATAGAAGAGTCTGTTGTTCCACTCTACTAGAGGACATAC 778
Qy      276  GlyValGlnAanHisTyrLeuGly 283
Db      779  GGCCACAGCAACATCTTTGGTGGT 802

RESULT 3
CB025968
LOCUS
DEFINITION
  TgESTzyd32a10.y1 TGRH Tachyzoite Norm 7 cDNA Library Toxoplasma
  gondii cDNA clone TgESTzyd32a10.y1 5' similar to TR:P90613 P90613
  LACTATE DEHYDROGENASE ;, mRNA sequence.
ACCESSION
CB025968
VERSION
CB025968.1 GI:27722340
KEYWORDS
EST.
SOURCE
Toxoplasma gondii
ORGANISM
Toxoplasma gondii
Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
Sarcocystidae; Toxoplasma.
REFERENCE
1 (bases 1 to 784)
  Tang,K., Cole,R., Fogarty,S., Sibley,L.D., Ajioke,J.A., White,M.,
  Clifton,S., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M.,
  Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M.,
  Ritter,E., Bennett,J., Franklin,C., Tsagaris,I., Ronko,I.,
  Kennedy,S., Maguire,L., Waterston,R. and Wilson,R.
  Toxoplasma EST Project
  Unpublished (2001)
  Contact: Clifton, S.
  Toxoplasma EST Project
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: toxo@wustl.edu
  Contact David Sibley (toxost@borcim.wustl.edu) for further
  information relating to organism, libraries, or clone availability.
  Seq primer: -40RP from Gibco
  High quality sequence stop: 440.
  Location/Qualifiers
    1..784
      /organism="Toxoplasma gondii"
      /mol_type="mRNA"
      /strain="RH (Type 1)"
      /db_xref="taxon:5811"

FEATURES
source

```

```

/clone="TgESTzyd32a10.y1"
/dev_stage="Tachyzoite"
/lab_host="DH10B (GeneHog, Invitrogen, Inc.)"
/clone_lib="TGRH Tachyzoite Norm 7 cDNA Library"
/notes="Vector: pBluscript SK-; Site 1: EcoRI; Site 2:
XhoI; Toxoplasma RH strain tachyzoites were grown in human
foreskin fibroblast cultures in vitro. The library was
originally constructed by K.L.Wan, Cambridge University.
cDNAs were synthesized from polyA RNAs by oligo d(T)
priming and directionally cloned into the EcoRI to XhoI
sites of the Lambda ZapII vector using the ZAP-cDNA
synthesis kit (Stratagene). The primary cDNA library was
mass excised as phagemid using EXassist helper phage
(Scratagene). Phagemid DNA was extracted by
phenol-chloroform method, and hybridized against a pool of
highly abundant genes which were derived from short-cycle
PCR of the primary cDNA library. The normalized library
was electroporated into DH10B (GeneHog, Invitrogen, Inc).
WARNING: the library contains a small percentage of cDNAs
derived from the human host cells."

```

ORIGIN

```

Alignment Scores:
Pred. No.:      2,47e-76      Length:      784
Score:          742.50      Matches:    140
Percent Similarity: 77.3%      Conservatives: 34
Best Local Similarity: 62.2%      Mismatches: 50
Query Match:    44.0%      Indels:     1
DB:             6          Gaps:         1

US-09-390-846-2 (1-330) x CB025968 (1-784)

Qy      9  ArgProLysIleAlaMetValGlySerGlyMetIleGlyGlyThrMetAlaPheLeuCys 28
Db      110 AGAAGAGAGGTGGCCATGATTGGCTCTGGCATGTTGTGGCACTATGGCTACCTGCTGC 169
Qy      29  SerLeuArgGluLeuGlyAapValValLeuPheAapValValProMetGly 48
Db      170 GCTCTCCGTGAGTCGTCGTCAGCTGTTCTACGATGTTGTCAAGAGTATGCCGAGGT 229
Qy      49  LysAlaMetAapIleSerHisAanSerSerValValAapThrGlyIleThrValTyrGly 68
Db      230 AAGGCTCTTGACCTGAGCCATGACCTCGCTGGTGCACCAACAGTTTCCTCGGTGCT 289
Qy      69  SerAanSerTyrGlu---CysLeuLysGlyAlaAapValValIleIleThrAlaGlyIle 87
Db      290 GAGTACTCTTACGAGCGCGCTCACCGGTGGAGTATGCTTATCGTACCGCGGTCTG 349
Qy      88  ThrLysIleProGlyLysSerAapLysGluTyrSerArgMetAapLeuProValAan 107
Db      350 ACCAAGTGTGGGGCAAGCCGACTCCGAGTGGAGCGCAACGATCTGCTCCGTTCAAC 409
Qy      108 IleLysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAanAlaPheVal 127
Db      410 TCGAAGATCATTCGCGAGATCGGTGAGACATCAGAAGTACTGCCCAAGACCTTCATC 469
Qy      128 IleAanIleThrAanProLeuAapValMetValAlaAlaLeuGlnGluSerSerGlyLeu 147
Db      470 ATCGTGGTACCACACCGCTGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 529
Qy      148 ProHisHisArgIleCysGlyMetAlaGlyMetLeuAapSerSerArgPheArgMet 167
Db      530 CGGACCAACATGATCTCGGTATGGCTTGCATGCTCGACTCTGCTGCTGCTGCTGCTG 599
Qy      168 IleAlaAapLysLeuGluValSerProArgAapValGlnGlyMetValIleGlyValHis 187
Db      590 GTCCGCGAGCGCTCTCTGCTCTCTCCCGGACGTCAGGCCACCGCTCATCGCACAC 649
Qy      188 GlyAapHisMetValProLeuSerArgTyrAlaThrValAanGlyIleProLeuSerGlu 207
Db      650 GCGACTGTCATGTCCTCGCTTGTCCGGTACATTACCGTGAACGGCTACCGGATCCAGAG 709
Qy      208 PheValLysLysGlyTyrIleLysGlnGluValAapIleValGlnLysThrLys 227

```



```

/cb_xref="taxon:5811"
/clone="T9ESTryc77h03.y1"
/dev_sage="tachyzoite"
/lab_host="DH10B (GeneHog, Invitrogen, Inc.)"
/clone_lib="T9RH Tachyzoite Norm 5 cDNA Library"
/notes="Vector: pBluscript SK-; Site 1: EcoRI; Site 2:
XhoI; Toxoplasma RH strain tachyzoites were grown in human
foreskin fibroblast cultures in vitro. The library was
originally constructed by K.L.Wan, Cambridge University.
cDNAs were synthesized from polyA RNAs by oligo d(T)
priming and directionally cloned into the EcoRI to XhoI
sites of the Lambda ZapII vector using the ZAP-cDNA
synthesis kit (Stratagene). The primary cDNA library was
mass excised as phagemid using EXassist helper phage
(Stratagene). Phagemid DNA was extracted by
phenol-chloroform method, and hybridized against a pool of
highly abundant genes which were derived from short-cycle
PCR of the primary cDNA library. The normalized library
was electroporated into DH10B (GeneHog, Invitrogen, Inc).
WARNING: the library contains a small percentage of cDNAs
derived from the human host cells."

ORIGIN
Alignment Scores:
Pred. No.: 6.9e-74 Length: 753
Score: 721.50 Matches: 137
Percent Similarity: 76.7% Conservative: 34
Best Local Similarity: 61.4% Mismatches: 51
Query Match: 42.7% Indels: 1
DB: 6 Gaps: 1

US-09-390-846-2 (1-330) x CB025472 (1-753)

QY 9 ArgProlylAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
DB 85 AGAAGAAGGTGGCCATGATGGCTCTGGCATGATTGGTGGCATTGGCTACCTGTGC 144

QY 29 SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMetGly 48
DB 145 GCTCTCGGTGAGTGGCTGACGTCTCTACGATGTTGTCAAGAGGTATGCCGAGGTT 204

QY 49 LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyrGly 68
DB 205 AAGCTCTTACCTTGACCCAGTACGCTCCGTGGTGCACACCAACAGTTTCCTCGGTGCT 264

QY 69 SerAsnSerTyrGlu---CysLeuLysGlyAlaAspValValIleIleThrAlaGlyIle 87
DB 265 GAGTACTTTCAGAGTGGCGCTCACCCTGGCGGACTGCGTTATCGTTACCGCGGTCTG 324

QY 88 ThrLysIleProGlyLysSerAspLysGluTyrSerArgMetAspLeuLeuProValAsn 107
DB 325 ACCAAGTGGCGGCAAGCCGACTCCGAGTGGAGCGCAACGATCTGCTCCCGTTCAAC 384

QY 108 IleIysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAlaPheVal 127
DB 385 TCAGAGATCAATTCGCGAGATCGGTGAGAACATCAAGAGTACTGCCCAAGACCTTCATC 444

QY 128 IleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGluSerSerGlyLeu 147
DB 445 ATGTGTTCAACCAACCCGCTGGACTGCTGATGTTAAGTTATGCGGAGGCTCTGGCGTC 504

QY 148 ProHisGlyArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMet 167
DB 505 CCACCAACATGATCTCGGTATGGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 564

QY 168 IleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHis 187
DB 565 GTGCCGACGCGTCTCTGCTCTCTCCCGGAGCTCCAGGCCACCGCTCATCGGCACAC 624

QY 188 GlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGlu 207
DB 625 GGCAGTGCATGTTGTCGCGGACATTACCGTGAACCGCTACCGCATCCAGAG 684

```

208 PheValLysGlyThrIleLysGlnGluValAspAspIleValGlnLysThrLys 227
685 TTCTCAAGGACGCGTAGTACGAGAGACGCTCGAGGAGTCTGAGCACCCAAA 744

228 ValAlaGly 230
745 GTGTCTGGC 753

RESULT 6
CX022234
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CX022234 900 bp mRNA linear EST 01-JAN-2005
EST940553 Field isolate cDNA library Plasmodium vivax cDNA clone
PVMJ731.5' end, mRNA sequence.
CX022234 GI:56961275
EST.
Plasmodium vivax (malaria parasite P. vivax)
Plasmodium vivax
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
Carlton, J.M. and Cui, L.
A survey of genes in Plasmodium vivax by EST sequencing
Unpublished (2004)
Contact: Jane Carlton
Parasite Genomics Group
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-530-9319
Fax: 301-838-0208
Email: carlton@tigr.org
Seq primer: T1.
Location/Qualifiers
1..900
/organism="Plasmodium vivax"
/mol_type="mRNA"
/strain="Field isolate"
/db_xref="taxon:5855"
/clone="PVMJ731"
/notes="Field isolate cDNA library"
Site 1: Sfi 1A; Site 2:
Site 1B: Plasmodium vivax field isolate cDNA library made
in lambda Triplex2. Inserts cloned unidirectionally in the
Sfi 1A and Sfi 1B sites. Mass excision of library produced
inserts in pTriplex2 plasmid. Inserts sequenced from
either 5' or 3' end using Triplex2 sequencing primer or
polydT 24 bp primer respectively."

Alignment Scores:
Pred. No.: 2.33e-73 Length: 900
Score: 718.00 Matches: 144
Percent Similarity: 70.9% Conservative: 39
Best Local Similarity: 55.8% Mismatches: 73
Query Match: 42.5% Indels: 2
DB: 8 Gaps: 1

US-09-390-846-2 (1-330) x CX022234 (1-900)

QY 66 ValTyrGlySerAsnSerTyrGluCysLeuLysGlyAlaAspValValIleIleThrAla 85
DB 31 GTGACTGGCTCGAAGCTCGTATGATGATCTGTAAGGGAGCGGACGTGGTGTGCTACCTGC 90

QY 86 GlyIleThrLysIleProGlyLysSerAspLysGluTyrSerArgMetAspLeuPro 105
DB 91 GGNATTACTTAAGCACCAAGAAAGACGACAGGAATGGAACCGATGATTTACTCTCCC 150

QY 106 ValAsnIleLysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAla 125
DB 151 TTGAATAACAAATATTGTTGAGATTGGGGACATATTAAAGAACTTTGGCCCAATGCC 210

QY 126 PheValIleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnSerSer 145
DB 211 TTATCATTTGTGTGACGAACCCAGTGGAGTGTGATGTTACTTCTTCGAGCATTC 270


```

Db      661 GAGGAGGAAAGAGCTTCCAA 684
      ... ..
      772 bp mRNA linear EST 13-JAN-2003
      TgESTzyc99b11.y1 TgrH Tachyzoite Norm 5 cDNA Library Toxoplasma
      gondii cDNA clone TgESTzyc99b11.y1 5' similar to TR:P90613 P90613
      LACTATE DEHYDROGENASE ; mRNA sequence.
      CB022051
      CB022051.1 GI:27692163
      EST.
      Toxoplasma gondii
      Toxoplasma gondii
      Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
      Sarcocystidae; Toxoplasma.
      1 (bases 1 to 772)
      Tang,K., Cole,R., Fogarty,S., Sibley,L.D., Ajioke,J.A., White,M.,
      Clifton,S., Pape,D., Martin,J., Wyllie,T., Dante,M., Marra,M.,
      Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M.,
      Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R., Ronko,I.,
      Kennedy,S., Maguire,L., Waterston,R. and Wilson,R.
      Toxoplasma EST Project
      Unpublished (2001)
      Contact: Clifton, S.
      Toxoplasma EST Project
      Washington University School of Medicine
      4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
      Tel: 314 286 1800
      Fax: 314 286 1810
      Email: toxo@watson.wustl.edu
      Contact David Sibley (toxost@borcim.wustl.edu) for further
      information relating to organism, libraries, or clone availability.
      Seq primer: -40RP from Gibco
      High quality sequence stop: 442.
      FEATURES
      Location/Qualifiers
      1..772
      /organism="Toxoplasma gondii"
      /mol_type="mRNA"
      /strain="RH (Type I)"
      /db_xref="taxon:5811"
      /clone="TgESTzyc99b11.y1"
      /dev_stage="Tachyzoite"
      /lab_host="DH10B (GeneHog, Invitrogen, Inc.)"
      /clone_lib="TgrH Tachyzoite Norm 5 cDNA Library"
      /notes="Vector: pBluscript SK-; Site 1: EcoRI; Site 2:
      XhoI; Toxoplasma RH strain tachyzoites were grown in human
      foreskin fibroblast cultures in vitro. The library was
      originally constructed by K.L.Wan, Cambridge University.
      cDNAs were synthesized from polyA RNAs by oligo d(T)
      priming and directionally cloned into the EcoRI to XhoI
      sites of the Lambda ZapII vector using the ZAP-cDNA
      synthesis kit (Stratagene). The primary cDNA library was
      mass excised as phagemid using EXAssist helper phage
      (Stratagene). Phagemid DNA was extracted by
      phenol-chloroform method, and hybridized against a pool of
      highly abundant genes which were derived from short-cycle
      PCR of the primary cDNA library. The normalized library
      was electroporated into DH10B (GeneHog, Invitrogen, Inc).
      WARNING: the library contains a small percentage of cDNAs
      derived from the human host cells."
      ORIGIN
      Alignment Scores:
      Pred. No.: 9,47e-70 Length: 772
      Score: 686.50 Matches: 131
      Percent Similarity: 77.2% Conservative: 25
      Best Local Similarity: 64.9% Mismatches: 45
      Query Match: 40.7% Indels: 1
      DB: 6 Gaps: 1
      US-09-390-846-2 (1-330) x CB022051 (1-772)
  
```

```

QY      9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyGlyThrMetAlaPheLeuCys 28
      |||
      167 AGAAGAAGAGTGGCCATGATGGCTCTGGCATGATGGTGGCCTATGGCTACCTGTGC 226
      |||
      29 SerLeuArgGluLeuGlyAspValValLeuPheAspValValProMetGly 48
      |||
      227 GCTCTCCGTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 286
      |||
      49 LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyGly 68
      |||
      287 AAGGCTCTTGGACCTGAGCCATGACCTCGGTGGTGCACCAACGTTTCCGTCGGTCTG 346
      |||
      69 SerAsnSerTyGlu---CysLeuLysGlyAlaAspValValIleIleThrAlaGlyIle 87
      |||
      347 GAGTACTCTTACGAGGCCGCGCTCACCGTGGGACTGCTTATCGTTATCCGCGGCTCTG 406
      |||
      88 ThrLysIleProGlyLysSerAspLysGluTTPSerArgMetAspLeuLeuProValAsn 107
      |||
      407 ACCAAGTCCGGGCAAGCCGACTCCGAGTGGAGCCGAAACGATCTGCTCCGTTCAAC 466
      |||
      108 IleLysIleMetArgGluValGlyAlaAlaIleLysSerTyCysProAsnAlaPheVal 127
      |||
      467 TCGAGATCATTCGCGAGATCGGTGAGACATCAAGACTACTGCCCAAGACCTTCATC 526
      |||
      128 IleAsnIleThrAsnProLeuAspValMetValAlaLeuGlnGluSerSerGlyLeu 147
      |||
      527 ATCGTGGTCAACCAACCCGCTGGACTGCATGGTCAAGTTATATGCGAGGCTCTGGCGTC 586
      |||
      148 ProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMet 167
      |||
      587 CCGACCAACATGATCTGCGGTATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 646
      |||
      168 IleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHis 187
      |||
      647 GTCCGCGACGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 706
      |||
      188 GlyAspHisMetValProLeuSerArgTyAlaThrValAsnGlyIleProLeuSerGlu 207
      |||
      707 GCGGACTGCATGCTCCGCTTGTCCGTTACATTACCGTGAACGCTACCCGATCCAGAG 766
      |||
      208 PheVal 209
      |||
      767 TTCATC 772
  
```

CB030949 780 bp mRNA linear EST 13-JAN-2003
 TgESTzyc46c08.y1 TgrH Tachyzoite Norm 7 cDNA Library Toxoplasma
 gondii cDNA clone TgESTzyc46c08.y1 5' similar to TR:P90613 P90613
 LACTATE DEHYDROGENASE ; mRNA sequence.
 CB030949
 CB030949.1 GI:27727321
 EST.
 Toxoplasma gondii
 Toxoplasma gondii
 Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
 Sarcocystidae; Toxoplasma.
 1 (bases 1 to 780)
 Tang,K., Cole,R., Fogarty,S., Sibley,L.D., Ajioke,J.A., White,M.,
 Clifton,S., Pape,D., Martin,J., Wyllie,T., Dante,M., Marra,M.,
 Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M.,
 Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R., Ronko,I.,
 Kennedy,S., Maguire,L., Waterston,R. and Wilson,R.
 Toxoplasma EST Project
 Unpublished (2001)
 Contact: Clifton, S.
 Toxoplasma EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: toxo@watson.wustl.edu
 Contact David Sibley (toxost@borcim.wustl.edu) for further

information relating to organism, libraries, or clone availability.

Seq primer: -40RP from Gibco
High quality sequence stop: 436.

FEATURES

source
Location/Qualifiers
1..780
/organism="Toxoplasma gondii"
/mol_type="mRNA"
/strain="RH (Type I)"
/db_xref="taxon:5811"
/dev_stage="Tachyzoite"
/lab_host="DH10B (GeneHog, Invitrogen, Inc.)"
/clone_lib="TGRH Tachyzoite Norm 7 cDNA Library"
/note="Vector: pBluscript SK-; Site 1: EcoRI; Site 2: XhoI; Toxoplasma RH strain tachyzoites were grown in human foreskin fibroblast cultures in vitro. The library was originally constructed by K.L.Wan, Cambridge University. cDNAs were synthesized from polyA RNAs by oligo d(T) priming and directionally cloned into the EcoRI to XhoI sites of the Lambda ZapII vector using the ZAP-cDNA synthesis kit (Stratagene). The primary cDNA library was mass excised as phagemid using ExAssist helper phage (Stratagene). Phagemid DNA was extracted by phenol-chloroform method, and hybridized against a pool of highly abundant genes which were derived from short-cycle PCR of the primary cDNA library. The normalized library was electroporated into DH10B (GeneHog, Invitrogen, Inc). WARNING: the library contains a small percentage of cDNAs derived from the human host cells."

ORIGIN

Alignment Scores:
Pred. No.: 2,17e-69 Length: 780
Score: 683.50 Matches: 131
Percent Similarity: 77.1% Conservative: 24
Best Local Similarity: 65.2% Mismatches: 45
Query Match: 40.5% Indels: 1
DB: 6 Gaps: 1

US-09-390-846-2 (1-330) x CB030949 (1-780)

QY 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyGlyThrMetAlaPheLeuCys 28
DB 177 AGAAGAGGTGCCATGATGGCTTGGCATGATTGGTGGCACTATGGCTACCTGTGC 236
QY 29 SerLeuArgGluLeuGlyAspValValPheAspValValProAsnMetProMetGly 48
DB 237 GCTCTCCGTGAGCTCGTGCAGCTCGTCTCTACGATGTTGTCAAAGGTATGCCGAGGCT 296
QY 49 LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyrGly 68
DB 297 AAGGCTCTTGACCTGAGCCATGAGCTCCGTGGTGCACACCAACATCTCTCCGTCGTGT 356
QY 69 SerAsnSerTyrGlu---CysLeuLysGlyAlaAspValValIleIleThrAlaGlyIle 87
DB 357 GAGTACTTACAGGCGCGCTCACCGGTGGCGATCGGTATCTGTTACCGCGGCTCTG 416
QY 88 ThrLysIleProGlyLysSerAspLysGluTyrSerArgMetAspLeuLeuProValAsn 107
DB 417 ACCAAGGTGCGCGCAAGCCGACTCCGAGTGGAGCCGAACATCTGCTCCGTTCAAC 476
QY 108 IleLysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAlaPheVal 127
DB 477 TCGAAGATCATTCGCGAGATCGCTCAGAACATCAAGAAGTACTGCCCAAGACCTTCATC 536
QY 128 IleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeu 147
DB 537 ATCGTGGTCACCAACCGCTGGACTGCATGGTCAAGTTATGTGCCGAGGCTCTGCGCTC 596
QY 148 ProHisHieArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMet 167
DB 597 CCGACCAACATGATCTGCGGTATGGCTGCATGCTCGACTCTGCTGCTTCGCCCGCATAC 656

QY 168 IleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHis 187
DB 657 GTCCGCGAGCGCTCTCTGTCTCTCCCGGACGTCACAGGCCACCGTATCCGCACAC 716
QY 188 GlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGlu 207
DB 717 GCGGACTGCATGGTCCCGCTTGTCCGGTACATTACCGTGAACCGGTACCCGATCCAGAAG 776
QY 208 Phe 208
DB 777 TTC 779

RESULT 10

CB027630

LOCUS

CB027630

DEFINITION

gondii cDNA clone TgESTzyc98c08.y1 5' similar to TR:P90613 P90613

LACTATE DEHYDROGENASE ; mRNA sequence.

CB027630

CB027630.1 GI:27724002

EST.

Toxoplasma gondii

Toxoplasma gondii

Sarcocystidae; Toxoplasma.

1 (bases 1 to 721)

Tang,K., Cole,R., Fogarty,S., Sibley,L.D., Ajioke,J.A., White,M.,

Clifton,S., Pape,D., Martin,J., Wyllie,T., Dante,M., Marra,M.,

Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M.,

Ritter,E., Bennett,J., Franklin,C., Teagareishvili,R., Ronko,I.,

Kennedy,S., Maguire,L., Waterston,R. and Wilson,R.

Toxoplasma EST Project

Unpublished (2001)

Contact: Clifton, S.

Toxoplasma EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: toxo@watson.wustl.edu

Contact David Sibley (toxos@borcim.wustl.edu) for further

information relating to organism, libraries, or clone availability.

Seq primer: -40RP from Gibco

High quality sequence stop: 433.

Location/Qualifiers

1..721

/organism="Toxoplasma gondii"

/mol_type="mRNA"

/strain="RH (Type I)"

/db_xref="taxon:5811"

/clone="TgESTzyc98c08.y1"

/dev_stage="Tachyzoite"

/lab_host="DH10B (GeneHog, Invitrogen, Inc.)"

/clone_lib="TGRH Tachyzoite Norm 5 cDNA Library"

/note="Vector: pBluscript SK-; Site 1: EcoRI; Site 2: XhoI; Toxoplasma RH strain tachyzoites were grown in human foreskin fibroblast cultures in vitro. The library was originally constructed by K.L.Wan, Cambridge University. cDNAs were synthesized from polyA RNAs by oligo d(T) priming and directionally cloned into the EcoRI to XhoI sites of the Lambda ZapII vector using the ZAP-cDNA synthesis kit (Stratagene). The primary cDNA library was mass excised as phagemid using ExAssist helper phage (Stratagene). Phagemid DNA was extracted by phenol-chloroform method, and hybridized against a pool of highly abundant genes which were derived from short-cycle PCR of the primary cDNA library. The normalized library was electroporated into DH10B (GeneHog, Invitrogen, Inc). WARNING: the library contains a small percentage of cDNAs derived from the human host cells."

ORIGIN

Alignment Scores:

Pred. No.: 2.21e-68 Length: 721
 Score: 674.50 Matches: 130
 Percent Similarity: 77.2% Conservative: 22
 Best Local Similarity: 66.0% Mismatches: 44
 Query Match: 40.0% Indels: 1
 DB: 6 Gaps: 1

US-09-390-846-2 (1-330) x CB027630 (1-721)

QY 9 ArgProlystleAlaMetValGlySerGlyMetileGlyThrMetAlaPheLeuCys 28
 DB 131 AGAAAGAGTGGCCATGATGGCTCTGGCATGATTGGTGGCACTATGGCTACCTGTGC 190
 QY 29 SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMetGly 48
 DB 191 GCTCTCGTGGAGTGGCTGACGTGCTCTACGATGTTGTCAAAGGTATGCCCGAGGGT 250
 QY 49 LysAlaMetAspIleSerHisAsnSerValValAspThrGlyIleThrValTyrGly 68
 DB 251 AAGCTCTTACCTTACCTGACCATGTGACCTCCGTGGTCACACCAACGGTTTCCGTCGGTCT 310
 QY 69 SerAsnSerTyrGlu---CysLeuLysGlyAlaAspValValIleThrAlaGlyIle 87
 DB 311 GAGTACTCTTACGAGGCGCGCTCACCGTGGGACTGCGTATCTGTTACGCGCGTCTG 370
 QY 88 ThrLysIleProGlyLysSerAspLysGluTyrSerArgMetAspLeuLeuProValAsn 107
 DB 371 ACCAAGGTGCGGCAAGCCGACTCCGAGTGGAGCGAAACGATCTGCTCCGTTCAAC 430
 QY 108 IleLysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAlaPheVal 127
 DB 431 TCGAAGATCATCTCGGAGATCGGTGAGAACATCAAGAACTGCTCCCAAGACCTTCATC 490
 QY 128 IleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeu 147
 DB 491 ATGCTGTCACCAACCGCTGGACTGCTGATGTCAGGTTATGCGGAGGCTCTGGCGTC 550
 QY 148 ProHisArgLysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMet 167
 DB 551 CCGACCAACATGATCTCGCGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 610
 QY 168 IleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHis 187
 DB 611 GTCGCGACCGCTCTCTGCTCTCCCGCGAGTCCAGGCCACCGCTCATCGGCACAC 670
 QY 188 GlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIlePro 204
 DB 671 GCGGACTGCTGCTCCGCTTGTCCGTTATACCTGACCGTACCGG 721

RESULT 11
 CF259970
 LOCUS
 DEFINITION NC0391h09.y2 NC-LIV Tachyzoite cDNA Library Neospora caninum
 cDNA 5' similar to TR:P90613 P90613 LACTATE DEHYDROGENASE ;, mRNA
 sequence.
 ACCESSION CF259970
 VERSION
 KEYWORDS EST.
 SOURCE
 ORGANISM Neospora caninum
 Neospora caninum
 Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
 Sarcocystidae; Neospora.
 REFERENCE 1 (bases 1 to 630)
 AUTHORS Cole, R., Fogarty, S., Tang, K., Howe, D.K., Sibley, L.D., Clifton, S.,
 Marra, M., Hillier, L., Pape, D., Martin, J., Wyllie, T., Theising, B.,
 Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Ronko, I.,
 Teagareishvili, R., Fedele, M., Belaygorod, L., Franklin, C.,
 Carr, L.M., Grw, A., Maguire, L., Wadkins, J., Richey, J., Waterston, R.
 and Wilson, R.
 TITLE USDA-WashU Neospora EST Project
 JOURNAL Unpublished (2000)
 COMMENT Contact: Sandy Clifton, Ph.D. - Neospora
 USDA-WashU Neospora EST Project

FEATURES
 source
 1. 630
 /organism="Neospora caninum"
 /mol_type="mRNA"
 /db_xref="taxon:29176"
 /dev_stage="Tachyzoite"
 /lab_host="Nc-LIV Tachyzoite cDNA Library"
 /clone_lib="Nc-LIV Tachyzoite cDNA Library"
 /notes="Vector: pBluescript II SK+ vector type: plasmid;
 Site 1: EcoRI; Site 2: XhoI; The cDNA library was
 constructed by Kellang Tang, and Robert Cole at Washington
 University. cDNA was synthesized from Poly(A)+ mRNA using
 an oligo-d(T) primer containing a XhoI site. Following
 second strand synthesis, EcoRI adapters were ligated to
 the cDNA, and products were size-selected on sephacryl
 S500. The cDNA were directionally cloned into the
 EcoRI/XhoI prepared pBluescript II SK+ vector, and
 electroporated into Electropen Blue cells (Stratagene).
 The library may contain a small percentage of host or
 bacterial contaminants."

ORIGIN
 Alignment Scores:
 Pred. No.: 4.07e-68 Length: 630
 Score: 671.50 Matches: 128
 Percent Similarity: 76.1% Conservative: 31
 Best Local Similarity: 61.2% Mismatches: 49
 Query Match: 39.8% Indels: 1
 DB: 6 Gaps: 1

US-09-390-846-2 (1-330) x CF259970 (1-630)

QY 32 GluLeuGlyAspValValLeuPheAspValProAsnMetProMetGlyLysAlaMet 51
 DB 3 GAGTTGGCTGACGTGCTCTCTACGATGTTGTCAAAGGCATGCCGAGGCAAGGGCTG 62
 QY 52 AspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyrGlySerAsnSer 71
 DB 63 GATCTCAGCCACGTGACGTCCGTCGTCGACACGACGTGCTGCTGCGGACACTCG 122
 QY 72 TyrGlu---CysLeuLysGlyAlaAspValValIleThrAlaGlyIleThrLysIle 90
 DB 123 TAGCAAGCCGCTCACCGGTGGGACTGCTTATTGTCACCGCGGTCTGACCAAGGTT 182
 QY 91 ProGlyLysSerAspLysGluTyrSerArgMetAspLeuProValAsnIleLysIle 110
 DB 183 CCGGGCAAAACCCGACTCGGAGTGGAGCCGACGACCTCTGCTGCTTCAACTCCAAGATC 242
 QY 111 MetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAlaPheValIleAsnIle 130
 DB 243 ATCCGCGAGATCGGCCAGAACATCAAGAGTACTGCCAAGACCTTCATCATGTGGTG 302
 QY 131 ThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeuProHisHis 150
 DB 303 ACGAACCTCTGGACTGATGTTGTCAAAGTCATGCTGGAGGCTCCGCGCTCCCAAGAAC 362
 QY 151 ArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMetIleAlaAsp 170
 DB 363 ATGATCTGGGTATGGCTTGCATGCTGCGACTCCGCGCGCTTCGCGCGCTACGTCGTGAC 422
 QY 171 LysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHisGlyAspHis 190
 DB 423 GCCTCTGCTGCTCTCTCTCGGACGTCACAGGCCACATCATCGGACGACGCGGACTGC 482
 QY 191 MetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGluPheValLys 210

Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Contact David Sibley (toxoeat@borcim.wustl.edu) for further
 information relating to organism, libraries, or clone availability.
 Seq primer: T3.

information relating to organism, libraries, or clone availability.

Seq primer: -40RP from Gibco
High quality sequence stop: 567.

FEATURES

source

1. 628
/organism="Toxoplasma gondii"

/mol_type="mRNA"

/strain="tachyzoite"

/db_xref="taxon:5811"

/clones="TGSTzym82h11.y1"

/dev_stage="tachyzoite"

/lab_host="ElectroTen Blue cells (Stratagene)"

/clone_lib="TgMAS Tachyzoite cDNA Library"

/notes="Vector: pBluescript II SK⁺; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed by Kellang Tang,

and Robert Cole at Washington University. cDNA was

synthesized from Poly(A)⁺ mRNA using an oligo-d(T) primer

containing a XhoI site. Following second strand synthesis,

EcoRI adapters were ligated to the cDNA, and products were

size-selected on Sephacryl S500. The cDNA were

directionally cloned into the EcoRI/XhoI prepared

pBluescript II SK⁺ vector, and electroporated into

ElectroTen Blue cells (Stratagene). The library may

contain a small percentage of host or bacterial

contaminants."

ORIGIN

Alignment Scores:

Pred. No.: 1.57e-66 Length: 628
Score: 658.00 Matches: 122
Percent Similarity: 75.8% Conservations: 35
Best Local Similarity: 58.9% Mismatches: 50
Query Match: 39.0% Indels: 0
DB: 7 Gaps: 0

US-09-390-846-2 (1-330) x CN621209 (1-628)

Qy 95 AspLysGluTrpSerArgMetAspLeuLeuProValAsnLeuLeuMetArgGluVal 114
Db |||||
Qy 115 GlyAlaAlaLeuLysSerTyrCysProAsnAlaPheValLeuLeuAsnLeuLeuProLeu 134
Db |||||
Qy 62 GGTGAGAACATCAAGAGTACTGCCCCAAGACCTTCATCATCGTGGTCCACCAACCCCGCTG 121
Db |||||
Qy 135 AspValMetValAlaAlaLeuGlnGluSerSerGlyLeuProHisHisArgLeuCysGly 154
Db |||||
Qy 182 GACTGATGTCAAGTTATGTGCGAGGCTCTGGCGTACGTCGCGACCATGATCTCGGT 181
Db |||||
Qy 155 MetAlaGlyMetLeuAspSerSerArgPheArgMetLeuAlaAspLysLeuGluVal 174
Db |||||
Qy 175 SerProArgAspValGlnGlyMetValLeuGlnHisGlyAspHisMetValProLeu 194
Db |||||
Qy 242 TCTCCCGCGAGGCTCCGAGGCGCCGCTCATCGGACACACACGCGGACTCGATGGTCCCGCTT 301
Db |||||
Qy 195 SerArgTyrAlaThrValAsnGlyLeuProLeuSerGluPheValLysLysGlyTrpIle 214
Db |||||
Qy 302 GTCCGGTACATTACCGTGAACGGTACCGATCCAGAGTTCATCAAGAGCGGGTAGTC 361
Db |||||
Qy 215 LysGlnGluValAspAspIleValGlnLysThrLysValAlaGlyGlyGluLeuVal 234
Db |||||
Qy 362 ACGGAGAAGCAGCTCGAGGAGATCGCTGAGCACACCAAAAGTGTCTGCGCGGAGATCGTC 421
Db |||||
Qy 235 ArgLeuLeuGlnGlySerAlaTyrAlaProGlyAlaSerAlaLeuGlnMetAla 254
Db |||||
Qy 422 CGTCTCTCGGCCAGGGTTCGGCTTACTAGCCGCCCGCCCGCTGCTCGCATGGCA 481
Db |||||
Qy 255 GluSerTyrLeuLysAspArgLysArgValMetValCysSerCysTyrLeuGlnGlyGln 274
Db |||||
Qy 482 ACATCTCTTGTGAACGACGAAAGCGCGTATCCCGTGCAGTGTGTACTCAACGGAGAG 541
Db |||||

Qy 275 TyrGlyValGlnAsnHisTyrLeuGlyValProCysValIleGlyArgGlyValGlu 294
Db |||||
Qy 542 TAGCGTTGAAGGACATGTTTCATTGGTCCCGCGGCGTTCATTGGAGCGCCCGCCACGAG 601
Db |||||
Qy 295 LysIleLeuGluLeuGluLeu 301
Db |||||
Qy 602 CGGTCATCCAGTCCAGCTG 622
Db |||||

RESULT 14

CB025686

LOCUS

DEFINITION

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

TOXOPLASMA GONDII

ORIGIN

Alignment Scores:

Pred. No.: 3.05e-66 Length: 744
Score: 656.50 Matches: 127
Percent Similarity: 77.2% Conservations: 22
Best Local Similarity: 65.8% Mismatches: 43

FEATURES

Location/Qualifiers

1. 744

/organism="Toxoplasma gondii"

/mol_type="mRNA"

/strain="RH (Type I)"

/db_xref="taxon:5811"

/clone="TGESTzyc83a05.y1"

/dev_stage="tachyzoite"

/lab_host="DH10B (GeneHog, Invitrogen, Inc.)"

/clone_lib="TGR Tachyzoite Norm 5 cDNA Library"

/notes="Vector: pBluescript SK⁺; Site 1: EcoRI; Site 2: XhoI; Toxoplasma RH strain tachyzoites were grown in human foreskin fibroblast cultures in vitro. The library was originally constructed by K.L.Wan, Cambridge University. cDNAs were synthesized from polyA RNAs by oligo d(T) priming and directionally cloned into the EcoRI to XhoI sites of the Lambda ZapII vector using the ZAP-cDNA synthesis kit (Stratagene). The primary cDNA library was mass excised as phagemid using ExAssist helper phage (Stratagene). Phagemid DNA was extracted by phenol-chloroform method, and hybridized against a pool of highly abundant genes which were derived from short-cycle PCR of the primary cDNA library. The normalized library was electroporated into DH10B (GeneHog, Invitrogen, Inc). WARNING: the library contains a small percentage of cDNAs derived from the human host cells."

Db 708 GCGACTGCATGGTCCCGCTTGTCGGTACATTACCGTG 746

Search completed: March 3, 2006, 01:04:29
Job time : 3716 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 2, 2006, 23:42:12 ; Search time 1114 Seconds
(without alignments)
1974.281 Million cell updates/sec

Title: US-09-390-846-2

Perfect score: 1688

Sequence: 1 MAVFEKTRPKIAMVGSGL.....GSIDEVKEMQKAIADASK 330

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delpop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abss/ABSSWEB spool/US09390846/runat_02032006_104229_4815/app query.fasta.1
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs08
-USER=US09390846 @CGN.1.1.727 @runat_02032006_104229_4815 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOC=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 21:.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1685	99.8	1679	2 AAT51370	Elmeria 1
2	1034.5	61.3	1785	2 AAX91434	Aax91434 T. gondii
3	1034.5	61.3	1785	2 AAX91435	Aax91435 T. gondii
4	1034.5	61.3	1785	4 AAS42758	Aas42758 T. gondii

5	1034.5	61.3	1785	4 AAS42757	Aas42757 T. gondii
6	1034.5	61.3	1785	10 ADG17361	Adg17361 T. gondii
7	1034.5	61.3	1785	10 ADG17363	Adg17363 T. gondii
8	880.5	52.2	960	13 ADT43087	Adt43087 Bacterial
9	873.5	51.7	957	13 ADS59454	Ads59454 Bacterial
10	873.5	51.7	960	13 ADS62502	Ads62502 Bacterial
11	873.5	51.7	960	13 ADS62887	Ads62887 Bacterial
12	862.5	51.1	933	13 ADT45997	Adt45997 Bacterial
13	862	51.1	969	13 ADS60073	Ads60073 Bacterial
14	850.5	50.4	963	13 ADT42328	Adt42328 Bacterial
15	848.5	50.3	960	13 ADS656026	Ads656026 Bacterial
16	847.5	50.2	960	13 ADS58536	Ads58536 Bacterial
17	831.5	49.3	951	2 AAO72947	Aao72947 P. falcip
18	831.5	49.3	951	2 AAX26909	Aax26909 cDNA enco
19	740	43.8	939	13 ADT48486	Adt48486 Bacterial
20	709	42.0	945	13 ADT42644	Adt42644 Bacterial
21	706	41.8	1912	2 AAT17715	Aat17715 Heat resi
22	697	41.3	936	13 ADT41767	Adt41767 Bacterial
23	694	41.1	870	14 ACL69781	ACL69781 M. xanthu
24	694	41.1	1634	14 ACL63878	ACL63878 M. xanthu
25	689	40.8	855	13 ADT44376	Adt44376 Bacterial
26	675.5	40.0	959	13 ADS61598	Ads61598 Bacterial
27	672	39.8	927	13 ADS57058	Ads57058 Bacterial
28	652.5	38.7	912	13 ADS58414	Ads58414 Bacterial
29	649.5	38.5	978	13 ADT43469	Adt43469 Bacterial
30	641	38.0	978	6 ABN93233	Abn93233 Staphyloc
31	641	38.0	978	13 ADS04190	Ads04190 Staphyloc
32	635	37.6	912	13 ADS59185	Ads59185 Bacterial
33	635	37.6	951	13 ADT45099	Adt45099 Bacterial
34	634	37.6	975	13 ADS48079	Ads48079 Bacterial
35	632	37.4	5024	2 AAO55139	Aao55139 Staphyloc
36	632	37.4	5024	8 ABZ77355	Abz77355 Nucleotid
37	632	37.4	5024	10 AAL51843	Aal51843 Staphyloc
38	628	37.2	897	13 ADS57777	Ads57777 Bacterial
39	626.5	37.1	915	13 ADS58980	Ads58980 Bacterial
40	625.5	37.1	927	13 ADT44107	Adt44107 Bacterial
41	621	36.8	822	6 ABK72765	Abk72765 Bacillus
42	605.5	35.9	1008	13 ADS45395	Ads45395 Bacterial
43	569	33.7	1005	13 ADS45382	Ads45382 Bacterial
44	560.5	33.2	942	13 ADT48105	Adt48105 Bacterial
45	518	30.7	595	13 ADS62339	Ads62339 Bacterial

ALIGNMENTS

RESULT 1
AAT51370
ID AAT51370 standard; cDNA to mRNA; 1679 BP.
XX
AC AAT51370;
XX
XX
DT 17-OCT-2003 (revised)
DT 29-APR-1997 (first entry)
XX
DE Elmeria lactate dehydrogenase cDNA clone EASC2.
XX Lactate dehydrogenase; LDH; coccidiosis; vaccine; vector; ds.
XX Elmeria acervulina; strain Houghton.
XX
FH Key Location/Qualifiers
FT misc_RNA 1..61
FT /tag= b
FT /note= "pBluescriptII derived sequence"
FT CDS 280..1272
FT /tag= a
FT /product= "lactate dehydrogenase"
FT 1624..1679
FT /tag= c
FT /note= "pBluescriptII derived sequence"
XX
XX
PN AU9656287-A.
XX

immunogenic proteins and antibodies to the proteins can be used to inhibit *T. gondii* oocyst shedding in a cat due to infection with *T. gondii*. They can be used for preventing *T. gondii* infection and for preventing the spread of *T. gondii* infection. They can also be used for detecting *T. gondii* infection. The detection method can be used to detect parasite cysts or oocysts in feces, e.g. from enteric apicomplexa oocysts such as *Cryptosporidium* oocysts and *Toxoplasma* oocysts

SQ Sequence 1785 BP; 393 A; 490 C; 503 G; 399 T; 0 U; 0 Other;

[illegible]

US-09-390-846-2 (1-330) x AAX91434 (1-1785)

Qy	9	ArgProLysIleAlaMetValGlySerGlyMetIleGlyGlyThrMetAlaPheLeuCys	28
Db	109	AGAAAGAGGTGGCCATGATGGCTCTGGCATGATTGGTGGCAGCTATGGCTACCTGTGC	168
Qy	29	SerLeuArgGluLeuGlyAspValValIleuPheAspValValProAsnMetProMetGly	48
Db	169	GCTCTCCGTGGAGCTCGCTGACGCTGCTTCTACGATGTTGTCAAAGGATATGCCCGAGGCT	228
Qy	49	LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyrGly	68
Db	229	AAGGCTCTTGACCTGAGGCATGTGACCTCCGTGGTCGACCAACAGCTTTCCGTCCGTGCT	288
Qy	69	SerAsnSerTyrGlu---CysLeuLysGlyAlaAspValValIleIleThrAlaGlyIle	87
Db	289	GAGTACTCTTACGAGGCGCGCTACCGGTGGGACTGCGTTATGTTACCGCCCGGTCTG	348
Qy	88	ThrLysIleProGlyLysSerAspLysGluTyrPheArgMetAspLeuLeuProValAsn	107
Db	349	ACCAAGGTGCCGGCAAGCCGACTCCGAGTGGAGCCGAAACAGCATCTGCTCCGCTTCAAC	408
Qy	108	IleLysIleMetArgGluValGlyAlaIleLysSerTyrCysProAsnAlaPheVal	127
Db	409	TCGAAGATCATTCGGCAGATCGGTCAAGACATCAAGAAGTACTGCCCCAAGACCTTCATC	468
Qy	128	IleAsnIleThrAsnProLeuAspValMetValAlaLeuGlnGlnLysSerSerGlyLeu	147
Db	469	ATCGTGGTCACCAACCCGCTGGACTGCATGGTCAAGGTCAATGCGAGGCTCTGGCGTC	528
Qy	148	ProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgArgMet	167
Db	529	CCGACCAACATGATCTCGGGTATGGCCCTGCATGCTCGACTCTGGTCTCGCCGATAC	588
Qy	168	IleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHis	187
Db	589	GTCGCCGACGGCTGTCTCTCTCCCGGACGTCACAGGCCACCGTCATCGGCACACAC	648
Qy	188	GlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGlu	207
Db	649	GGCGACTGCATGGTCCCGCTTGTCCGGTACATTACCGTGACGACTACCCGATCCAGAAG	708
Qy	208	PheValLysLysGlyTyrIleLysGlnGluGluValAspAlleValGlnLysThrLys	227
Db	709	TTCTCAAGGACGGCGTAGTCACGGAGAAGCAGCTCGAGGAGATCGCTGAGCACACCAA	768
Qy	228	ValAlaGlyGlyIleValArgLeuGlnGlySerAlaTyrTyrAlaProGly	247
Db	769	GTGTCTGGCGCGAGATCGTCCGCTTCTCGGCCAGGGTTCGGCTTACTATGCGCCCGCC	828
Qy	248	AlaSerAlaIleGlnMetAlaGluSerTyrLysLysAspArgLysArgValMetValCys	267
Db	829	GCATCCCGTCTGGCCATGGCAACATCTCTTTGAAACGACGAAAAGCGGTCACTCCGTCG	888
Qy	268	SerCysTyrLeuGlnGlyIleTyrGlyValGlnAsnHisTyrLeuGluValProCysVal	287

QY 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyGlyThrMetAlaPheLeuCys 28

Db	889	AGTGTACTGCAACGGAGAGTAGCGGCTTGAAGGACATGTTCAATGCTGCCCGGCGGTC	948
Qy	288	IleGlyGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGluArgGln	307
Db	949	ATTGGAGCGCGCGCATCGAGCGCGTTCATCGAGCTCGAGCTGAACGAGGAGGAGAGAAG	1008
Qy	308	GluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaLeuAsp	327
Db	1009	CAGTTCACAGATCCGTCGAGCAGCTCATCGCGCTCAACAGCGGGTTGCTGCTTCAG	1068
Qy	328	Ala 328	
Db	1069	GCG 1071	
RESULT	3		
AAAX91435			
ID	AAAX91435	standard; DNA; 1785 BP.	

T. gondii MGIS6-5 reverse complement sequence.

KW Immunogenic protein; *Toxoplasma gondii* protein; oocyst shedding; cat;
 KW *T. gondii* infection; enteric apicomplexa oocyst; *Cryptosporidium* oocyst;
 KW *Toxoplasma* oocyst; ss.

05 Toxoplasma gondii.

AA PN WO9932633-A1.

01-JUL-1999.

18-DEC-1998: 98WO-US027137.

PR 19-DEC-1997; 97US-00994825.

PA (HESK-) HESKA CORP.

XX PI Milhausen MJ. Lutz SB. Na RK:

WPI; 1999-418930/35.

New isolated *Toxoplasma gondii* nucleic acids used, e.g. to treat PT infection caused by this microorganism.

XX
PS
Example 15: Page 342-344: 381pp: English.

The invention provides isolated *Toxoplasma gondii* nucleic acids that encode immunogenic polypeptides. The *T. gondii* nucleic acid molecules, immunogenic proteins and antibodies to the proteins can be used to inhibit *T. gondii* oocyst shedding in a cat due to infection with *T. gondii*. They can be used for preventing *T. gondii* infection and for preventing the spread of *T. gondii* infection. They can also be used for detecting *T. gondii* infection. The detection method can be used to detect parasite cysts or oocysts in feces, e.g. from enteric apicomplexa oocysts such as *Cryptosporidium* oocysts and *Toxoplasma gondii* oocysts.

Sequence 1785 BP; 393 A; 490 C; 503 G; 399 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.58e-101	Length:	1785
Score:	1034.50	Matches:	193
Percent Similarity:	77.3%	Conservative:	55
Best Local Similarity:	60.1%	Mismatches:	72
Query Match:	61.3%	Indels:	1
DB:	2	Gaps:	1

US-09-390-846-2 (1-330) x AAX91435 (1-1785)

```

Db 109 AGAAGAAGTGGCGCATGATTGGCTGGCATGTTGGTGGCATATGGGCTACCTGTC 168
Qy 29 SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMetGly 48
Db 169 GCTCTCGTAGCTCGCTGACGTGCTTCTTACCATGTTGTCAAAGGTATGCCGAGGCT 228
Qy 49 LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyGly 68
Db 229 AAGGCTCTTGACCTGAGCCATGTCACCTCCGTGGTGGACACCAACGTTTCGTCGCTGCT 288
Qy 69 SerAsnSerTyrGlu---CysLeuLysGlyAlaAspValValIleIleThrAlaGlyIle 87
Db 289 GAGTACTCTTACGAGCGCCGCTCACCCTGCGGACTCGCTTATCGTTACCGCCGGTCTG 348
Qy 88 ThrLysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProValAsn 107
Db 349 ACCAAGGTGCCGGGCAAGCCGACTCCGAGTGGAGCGCAACGATCTGCTCCGCTTCAAC 408
Qy 108 IleLysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAlaPheVal 127
Db 409 TCGAAGATCATTCGCGAGATCGGTGAGAACATCAAGAGTACTGCCCAAGACCTTCATC 468
Qy 128 IleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeu 147
Db 469 ATCGTGTGCACCAACCGCTGAGTCATGTCATGTCATGTCGAGGCTCTTGGCGTC 528
Qy 148 ProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMet 167
Db 529 CCGACCAACATGATCTGGCGGTATGGCTGTCATGTCGACTCTGGTTCGCTCCGCGATAC 588
Qy 168 IleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHis 187
Db 589 GTGCGGACGAGCGGTGTCGTCTCTCCCGCACGTCAGGACCGTCCATCGGCACACAC 648
Qy 188 GlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGlu 207
Db 649 GCGGACTCATGTCCTCCGCTTGTCCGGTACATTTACCGTGAACGACTCCGCTCAGAG 708
Qy 208 PheValLysLysGlyTrpIleLysGlnGluValAspIleValGlnLysThrLys 227
Db 709 TTCATCAAGGACGGCGTAGTCACGAGAGAGAGCTCGAGGAGATCGCTGAGCACACCAA 768
Qy 228 ValAlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaProGly 247
Db 769 GTGTCTGCGCGAGATCGTCCGCTCTCCGCGCAGAGGTTCGCTTACTACGCCCCCGCC 828
Qy 248 AlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCys 267
Db 829 GCATCCGCTGCGCCATGGCAACATCTCTTTGACGACGAAGAGCGGTCTATCCCGTGC 888
Qy 268 SerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysVal 287
Db 889 AGTGTGTACTGCAACGGAGAGTACGGCTTGAAGGACATGTTTCATTGGTCTCCCGCGCTC 948
Qy 288 IleGlyGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGluArgGln 307
Db 949 ATTGAGGCGCGGCGATCGAGCGGTCTATCGAGCTCGAGCTGAACGAGGAGAGAAGAAG 1008
Qy 308 GluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaAlaLeuAsp 327
Db 1009 CAGTCCAGAAAGTCCGTCGACGACGTATGGGCTCAACAGGCGGTTCGTCTTCAG 1068
Qy 328 Ala 328
Db 1069 GCG 1071

```

RESULT 4
 AAS42758
 ID AAS42758 standard; cDNA; 1785 BP.
 XX
 AC
 AAS42758;

DT 17-DEC-2001 (first entry)
 XX
 DE T. gondii cDNA encoding immunogenic protein PMGIS65.
 XX
 KW Immunogenic protein; oocyst; faeces; ss; enteric apicomplexa oocyst;
 KW Cryptosporidium oocyst; Toxoplasma oocyst; Giardia cyst; vaccine;
 KW oocyte shedding.
 XX
 OS Toxoplasma gondii.
 XX
 PN US2001014447-A1.
 XX
 PD 16-AUG-2001.
 XX
 PF 18-DEC-1998; 98US-00216393.
 XX
 PR 19-DEC-1997; 97US-00994825.
 XX
 PA (MILH/) MILHAUSEN M J.
 XX
 XX Milhausen MJ;
 XX
 XX WPI; 2001-529100/58.
 XX P-PSDB; AAU25549.
 DR
 DR
 XX
 XX
 PT Detecting parasite oocysts or cysts in feces, comprises eluting DNA from
 sample into aqueous solution by heating, amplifying DNA with primers
 PT specific for oocysts or cysts being detected, and detecting amplification
 product.
 XX
 XX
 PS Example 15; Page 165-166; 188pp; English.
 XX
 CC The invention relates to detection of parasite oocysts or cysts in a
 CC faeces sample comprising contacting the sample with a solid support,
 CC drying and then washing the sample with an aqueous wash solution, adding
 CC an aqueous elution solution and eluting DNA from the sample by heating
 CC or amplifying by PCR oocyst/cyst-specific DNA and detecting the
 CC amplification products. The method is useful for detecting parasite
 CC oocysts e.g., enteric apicomplexa oocysts such as Cryptosporidium oocysts
 CC or Toxoplasma oocysts, or for detecting parasite cysts e.g. Giardia
 CC cysts. The method is also useful for developing vaccines to prevent
 CC oocyte shedding in cats. The present sequence encodes an immunogenic
 CC protein from Toxoplasma gondii.
 XX
 SQ Sequence 1785 BP; 393 A; 490 C; 503 G; 399 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.58e-101 Length: 1785
 Score: 1034.50 Matches: 193
 Percent Similarity: 77.3% Conservative: 55
 Best Local Similarity: 60.1% Mismatches: 72
 Query Match: 61.3% Indels: 1
 DB: 4 Gaps: 1

US-09-390-846-2 (1-330) x AAS42758 (1-1785)

Qy 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
 Db 109 AGAAGAAGTGGCGCATGATTGGCTGGCATGTTGGTGGCATATGGGCTACCTGTC 168
 Qy 29 SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMetGly 48
 Db 169 GCTCTCGTAGCTCGCTGACGTGCTTCTTACCATGTTGTCAAAGGTATGCCGAGGCT 228
 Qy 49 LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyGly 68
 Db 229 AAGGCTCTTGACCTGAGCCATGTCACCTCCGTGGTGGACACCAACGTTTCGTCGCTGCT 288
 Qy 69 SerAsnSerTyrGlu---CysLeuLysGlyAlaAspValValIleIleThrAlaGlyIle 87
 Db 289 GAGTACTCTTACGAGCGCCGCTCACCCTGCGGACTCGCTTATCGTTACCGCCGGTCTG 348
 Qy 88 ThrLysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProValAsn 107
 Db 349 ACCAAGGTGCCGGGCAAGCCGACTCCGAGTGGAGCGCAACGATCTGCTCCGCTTCAAC 408
 Qy 108 IleLysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAlaPheVal 127
 Db 409 TCGAAGATCATTCGCGAGATCGGTGAGAACATCAAGAGTACTGCCCAAGACCTTCATC 468
 Qy 128 IleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeu 147
 Db 469 ATCGTGTGCACCAACCGCTGAGTCATGTCATGTCATGTCGAGGCTCTTGGCGTC 528
 Qy 148 ProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMet 167
 Db 529 CCGACCAACATGATCTGGCGGTATGGCTGTCATGTCGACTCTGGTTCGCTCCGCGATAC 588
 Qy 168 IleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHis 187
 Db 589 GTGCGGACGAGCGGTGTCGTCTCTCCCGCACGTCAGGACCGTCCATCGGCACACAC 648
 Qy 188 GlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGlu 207
 Db 649 GCGGACTCATGTCCTCCGCTTGTCCGGTACATTTACCGTGAACGACTCCGCTCAGAG 708
 Qy 208 PheValLysLysGlyTrpIleLysGlnGluValAspIleValGlnLysThrLys 227
 Db 709 TTCATCAAGGACGGCGTAGTCACGAGAGAGAGCTCGAGGAGATCGCTGAGCACACCAA 768
 Qy 228 ValAlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaProGly 247
 Db 769 GTGTCTGCGCGAGATCGTCCGCTCTCCGCGCAGAGGTTCGCTTACTACGCCCCCGCC 828
 Qy 248 AlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCys 267
 Db 829 GCATCCGCTGCGCCATGGCAACATCTCTTTGACGACGAAGAGCGGTCTATCCCGTGC 888
 Qy 268 SerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysVal 287
 Db 889 AGTGTGTACTGCAACGGAGAGTACGGCTTGAAGGACATGTTTCATTGGTCTCCCGCGCTC 948
 Qy 288 IleGlyGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGluArgGln 307
 Db 949 ATTGAGGCGCGGCGATCGAGCGGTCTATCGAGCTCGAGCTGAACGAGGAGAGAAGAAG 1008
 Qy 308 GluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaAlaLeuAsp 327
 Db 1009 CAGTCCAGAAAGTCCGTCGACGACGTATGGGCTCAACAGGCGGTTCGTCTTCAG 1068
 Qy 328 Ala 328
 Db 1069 GCG 1071

```

Db 349 ACCAAGTTCGGCGGCAAGCCGACTCCGAGTGGAGCCGAACGATCTGCTCCGGTTCAC 408
Qy 108 IleLysIleMetArgGluValGlyAlaIleLysSerTyrCysProAsnAlaPheVal 127
Db 409 TCGAAGATCAATTCGGCGAGATCGGTGAGAACATCAAGAGTAGTGTGCCCAAGACCTTCATC 468
Qy 128 IleAsnIleThrAsnProLeuAspValMetValAlaIleLeuGlnGluSerSerGlyLeu 147
Db 469 ATCGTGGTGACCAACCCGCTGAGCTGCGATGGTCAAGGTCAAGTGTGGAGCCCTTCGGCGTC 528
Qy 148 ProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMet 167
Db 529 CCGACCAACATCATCTCGGGTATGCGCTGCATCTCGACTTCGTGCTTCGCGCGATAC 588
Qy 168 IleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHis 187
Db 589 GTGCGGAGCGGCTGTCTCTCCCGGAGCGTCCAGGCCACCGTTCATCGGCACACAC 648
Qy 188 GlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyLeuProLeuSerGlu 207
Db 649 GGCAGCTCATGTTGTCCTGTCGCTGACATTCACCGTGAACGACTACCCGATCCAGAG 708
Qy 208 PheValLysLysGlyTyrIleLysGlnGluValAspAspIleValGlnLysThrLys 227
Db 709 TTCATCAGGACGGCTAGTCACGAGAGGAGCTCGAGGAGATCGCTGACACACCAAA 768
Qy 228 ValAlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaProGly 247
Db 769 GTGTCTGCGCGAGATCGTCCGCTTCCTCGGCCAGGGTTCGGCTTACTACGCCCGCC 828
Qy 248 AlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCys 267
Db 829 GCATCCGCTGTGCGCATGCGCAACATCTCTTGAACGACGAAAGCGCGTTCATCCCGTGC 888
Qy 268 SerCysTyrLeuGlnGlyGlyTyrValGlnAsnHisTyrLeuGlyValProCysVal 287
Db 889 AGTGTACTGCAACGGAGAGTACGGCTTGAGGACATGTTCAITGGTCTCCCGCGCGTC 948
Qy 288 IleGlyGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGluArgGln 307
Db 949 ATTGAGGCGCGGCATCGAGCGGTCTCATCGAGCTCGAGCTGAACGAGGAGAGAGAG 1008
Qy 308 GluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaLeuAsp 327
Db 1009 CAGTTCAGAGATCGTTCGACGACGTCATGGCGCTCAACAAAGCGGTGTGCTCTTCAG 1068
Qy 328 Ala 328
Db 1069 GCG 1071

RESULT 5
AAS42757
ID AAS42757 standard; cDNA, 1785 BP.
XX
AC AAS42757;
XX
DT 17-DEC-2001 (first entry)
XX
DE T. gondii cDNA encoding immunogenic protein PMGIS65.
XX
KW Immunogenic protein; oocyst; faeces; ss; enteric apicomplexa oocyst;
KW Cryptosporidium oocyst; Toxoplasma oocyst; Giardia cyst; vaccine;
KW oocyte shedding.
XX
OS Toxoplasma gondii.
XX
FN US2001014447-A1.
XX
PD 16-AUG-2001.
XX
PF 18-DEC-1998; 98US-00216393.
XX

```

```

PR 19-DEC-1997; 97US-00994825.
XX
PA (MILH/) MILHAUSEN M J.
XX
PI Milhausen MJ;
XX
DR WPI: 2001-529100/58.
DR P-FSDB; AAU25549.
XX
PT Detecting parasite oocysts or cysts in feces, comprises eluting DNA from
PT sample into aqueous solution by heating, amplifying DNA with primers
PT specific for oocysts or cysts being detected, and detecting amplification
PT product.
XX
PS Example 15; Page 164-165; 189pp; English.
XX
CC The invention relates to detection of parasite oocysts or cysts in a
CC faeces sample comprising contacting the sample with a solid support,
CC drying and then washing the sample with an aqueous wash solution, adding
CC an aqueous elution solution and eluting DNA from the sample by heating
CC and amplifying by PCR oocyst/cyst-specific DNA and detecting the
CC amplification products. The method is useful for detecting parasite
CC oocysts e.g., enteric apicomplexa oocysts such as Cryptosporidium oocysts
CC or Toxoplasma oocysts, or for detecting parasite cysts e.g. Giardia
CC cysts. The method is also useful for developing vaccines to prevent
CC oocyte shedding in cats. The present sequence encodes an immunogenic
CC protein from Toxoplasma gondii
XX
SQ Sequence 1785 BP; 393 A; 490 C; 503 G; 399 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.58e-101 Length: 1785
Score: 1034.50 Matches: 193
Percent Similarity: 77.3% Conservative: 55
Best Local Similarity: 60.1% Mismatches: 72
Query Match: 61.3% Indels: 1
DB: 4 Gaps: 1

US-09-390-846-2 (1-330) x AAS42757 (1-1785)
Qy 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheIleuCys 28
Db 109 AGAAAGAGGTGGCCATGATTTGGCTCTGGCATGATTTGGTGCATCTATGGGTACTGTGTC 168
Qy 29 SerLeuArgGluLeuGlyAspValValIlePheAspValValProAsnMetProMetGly 48
Db 169 GCTTCCTCGTGGCTCGCTGAGTGTCTCTACATGTTGTCAAGGTATGCCCGAGGT 228
Qy 49 LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyrGly 68
Db 229 AAGGCTCTTGACCTTGAGCCATGTGACCTCGGTGTCGACACCAACGTTTCGGTCCGTGCT 288
Qy 69 SerAsnSerTyrGlu---CysLeuLysGlyAlaAspValValIleIleThrAlaGlyIle 87
Db 289 GAGTACTCTTACGAGGCGCGCTCACCGGTGCGGACTGCGGTATCTGTTACCGCGGTCTG 348
Qy 88 ThrLysIleProGlyLysSerAspLysGluTyrSerArgMetAspLeuLeuProValAsn 107
Db 349 ACCNAGTGGCGGCAAGCCGAGCTCCGAGTGGAGCCGAACAGATCTGCTCCCTTCAC 408
Qy 108 IleLysIleMetArgGluValGlyAlaIleLysSerTyrCysProAsnAlaPheVal 127
Db 409 TCGAAGATCAATTCGGCGAGATCGGTGAGAACATCAAGAGTAGTGTGCCCAAGACCTTCATC 468
Qy 128 IleAsnIleThrAsnProLeuAspValMetValAlaIleLeuGlnGluSerSerGlyLeu 147
Db 469 ATCGTGGTGACCAACCCGCTGAGCTGCGATGGTCAAGGTCAAGTGTGTCGAGGCGCTTCGGCGTC 528
Qy 148 ProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMet 167
Db 529 CCGACCAACATGATCTCGGTGATGGCTGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 588
Qy 168 IleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHis 187

```

[illegible]

diseases caused by infection with T. gondii. The nucleic acid can be used as genetic vaccine which encodes the protein. The protein and the nucleic acid are used as diagnostic reagents for detection of T. gondii infection. The present sequence is used in the exemplification of the invention.

XX SQ Sequence 1785 BP; 393 A; 490 C; 503 G; 399 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.58e-101	Length:	1785
Score:	1034.50	Matches:	193
Percent Similarity:	77.3%	Conservative:	55
Best Local Similarity:	60.1%	Mismatches:	72
Query Match:	61.3%	Indels:	1
DB:	10	Gaps:	1

US-09-390-846-2 (1-330) x ADG17361 (1-1785)

Qy	9	ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys	28
Db	109	AGAAAGAAGGTGGCCATATTGGCTTCGCATGATTGGTGCACTATCGGCTACCTGTC	168
Qy	29	SerLeuArgGLuLeuGlyAspValValLeuPheAspValValProAsnMetPrometGly	48
Db	169	GCTCTCGGTGAGTCTGCGTAGCTCGTTCTTACCAGTGTGTCAAAGGTATGCCGAGG	228
Qy	49	LysAlaMetAepIleSerHisAsnSerValValAepThrGlyIleThrValTyrGly	68
Db	229	AAGGCTTTGNCTTGAGCCATGTCACTTCGTGTGACACCAAGTTTTCGTCGCTGCT	288
Qy	69	SerAenSertYrGlu--CysLeuLyseGlyAlaAepValValIleThrAlaGlyIle	87
Db	289	GAGTACTCTTACGAGGCGCGCTCACCGTTCGGACTCGTTATCGTTACCGCGGTCTG	348
Qy	88	ThrIlysileProGlyLysSerAspLysGluTrpSerArgMetAepLeuLeuProValAen	107
Db	349	ACCAAGGTGCGCGCAAGCCGACTCCGAGGTGGAGCCGAACGATCTGCTCCCGTTCAAC	408
Qy	108	IleLysileMetArgGluValcIcAlaAlaIleLysSertYrCysProAsnAlaPheVal	127
Db	409	TGGAAGATCATTCGGAGATCGGTTCAGAACATCAAGAAGTACTGCCCAAGACCTTCATC	468
Qy	128	IleAenIleThrAenProLeuAepValMetValAlaAlaLeuGlnGluSerSerglyLeu	147
Db	469	ATCGTGGTTCACAAACCCGCTGGACTGTCATGTGTCAAGGTCTATGTGCGAGCGCTCTGCGCTC	528
Qy	148	ProHishisArgIleCysGlyMetAlaClyMetLeuAepSerSerArgPheArgArgMet	167
Db	529	CCGACCAACATGATCTCGGGTATGGCTGTCATGTCGACTCTGTGCTTCGCCGATAC	588
Qy	168	IleAlaAepLysLeuGluValSerProArgAepValGlnGlyMetValIleGlyValHis	187
Db	589	GTCGCCGACGCGCTGTCTGTCTCTCCCCGCGACGTCAGGCCACCGTCATCGCACACAC	648
Qy	188	GlyAepHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGlu	207
Db	649	GGCGACTCATGGTCCCGCTTGTCCGGTACATTACCGTGAACGACTACCCGATCCAGAA	708
Qy	208	PheValIysLysGlyTrrIleLysGlnGluValAspApeIleValGlnLysThrLys	227
Db	709	TTTCAAGGACGCGGTAGTCAAGGAGGACGATCGAGGAGATCGCTGAGCACACCAAA	768
Qy	228	ValAlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlatYrTyrAlaProGly	247
Db	769	GTGTCTGCGCGGAGATCGTTCGCTTCCTCGGCCAGGGTTCGCTTACTACGCCCGCC	828
Qy	248	AlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCys	267
Db	829	GCATCCGCTGTGCGCATCGCAACATCCTCTTGAACGACGAAAAAGGGGTCTATCCGCTGC	888
Qy	268	SerCysTyrLeuGlnGlyClnTyrGlyValGlnAenHisTyrLeuGlyValProCysVal	287
Db	889	AGTGTGTACTCAACGAGAGATCGGTTTGAAGGACATGTTTCATTTGGTTCGCGCCGTC	948

KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polynucleotide; gene; ss.
 XX
 XX Bacteria.

OS US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAO/Y) CAO Y.

XX (HINK/) HINKLE G J.

XX (SLAT/) SLATER S C.

XX (CHEN/) CHEN X.

XX (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide
 XX for expression of a polynucleotide encoding a polypeptide from a
 XX microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 41525; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
 XX promoter functional in a plant cell, where the promoter is positioned to
 XX provide for expression of a polynucleotide encoding a polypeptide from a
 XX microbial source. The invention also relates to a transformed plant
 XX comprising the recombinant DNA construct and a method of producing a
 XX transformed plant having an improved property. The plant is a crop plant
 XX such as maize or soybean. The method of producing a transformed plant
 XX having an improved property comprises transforming a plant with the
 XX recombinant DNA construct and growing the transformed plant, where the
 XX polynucleotide or polypeptide is useful for producing plants with
 XX improved plant properties, e.g. improved cold, heat or drought tolerance,
 XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 XX increased resistance to plant disease, better growth rate by modification
 XX of the cell cycle pathway with plant growth regulators, increased rate of
 XX homologous recombination, modified seed oil or protein yield and/or
 XX content, improved yield by modification of carbohydrate, nitrogen or
 XX phosphorus use and/or uptake, by modification of photosynthesis or by
 XX providing improved plant growth and development under at least one stress
 XX condition, improved lignin production or improved galactomannan
 XX production. This sequence represents a bacterial polynucleotide used in
 XX the scope of the invention. Note: The sequence data for this patent did
 XX not form part of the printed specification but was obtained in electronic
 XX format from USPTO at seqdata.uspto.gov/sequence.html.

SQ Sequence 960 BP; 170 A; 325 C; 311 G; 154 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.06e-85	Length:	960
Score:	880.50	Matches:	173
Percent Similarity:	70.7%	Conservative:	54
Best Local Similarity:	53.9%	Mismatches:	89
Query Match:	52.2%	Indels:	5
DB:	13	Gaps:	1

US-09-390-846-2 (1-330) x ADT43087 (1-960)

Qy 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28

|||||

Db	7	CGCAAGAAGATCGCGCTGATCGGCTCGCGCATGATCGGTGGCACCTTCGGCACCTCGCT	66
Qy	29	SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMetGly	48
Db	67	GCAATCAAGAAATCGGCGATATCGTCTCTTGCATTTGCCAGAGGACGCCCGAGGCG	126
Qy	49	LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyGly	68
Db	127	AGGCGCTCGACATCGACAGTCCGGACCGGTGCAAGGCTTCGATGCAACCTCAAGGCG	186
Qy	69	SerAsnSerTyrGluCysLeuLysGlyAlaAspValValIleIleThrAlaGlyIleThr	88
Db	187	GCGAACAGCTACGAGGACATTGCGGGCGCGCATGCTGTCATCGTCACCGCGGTATCCCG	246
Qy	89	LysIleProGlyLysSerAspLysGluTyrSerArgMetAspLeuLeuProValAsnIle	108
Db	247	CGCAAGCGCGGC-----ATGAGCGCGACGATCTTCTCAAGACCAACCTG	291
Qy	109	LysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAlaPheValIle	128
Db	292	GGCGTGTATGAAGCCGTGGCGGAGGCGATGCGCGCGCACCGCCCGCGTTCGTGATC	351
Qy	129	AsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeuPro	148
Db	352	TGCATCACCAACCGCTCGACCGCATGCTGTGGCGCTGCGGAGTTCTTCGGGCTGCGG	411
Qy	149	HisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMetIle	168
Db	412	CACCAAGATCGTCGCGCATGCGCGGTGCTGCTGCGCGCGCTTCGCGCACTTCCTC	471
Qy	169	AlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHisGly	188
Db	472	GCCAGGAATTCACGCTCTCGTGCAGACGTCACCGCTTCGTGCTGGGCGGACACGCG	531
Qy	189	AspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGluPhe	208
Db	532	GACACCATGTCGCGGTGATCGAGTATTCGACCGTCGCGGCATCCCGTTCGCGACCTG	591
Qy	209	ValLysGlyTyrIleLysGlnGluValAspIleValGlnLysThrLysVal	228
Db	592	ATCAAGATGGCTGCTCCACCCAGGCGCATACGCCCATCTGTCGCGCGCACCGCTCG	651
Qy	229	AlaGlyGlyIleValArgLeuGlyGlnGlySerAlaTyrTyrAlaProGlyAla	248
Db	652	GCGCGCGGAGATCGTCGCTCTCAAGACCGCTCGGCTATTCGCGCGCACCC	711
Qy	249	SerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCysSer	268
Db	712	AGCGCGATCGCGATGGCGGAGAGCTATCTGAAGGACAAAGACGCTGCTTCCTGTGCC	771
Qy	269	CysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysValIle	288
Db	772	GCGCACCTCACCGGCCAGTACGCGCTCGACATCTGTACGTGCGCGTCCCATCGTCATC	831
Qy	289	GlyGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGluArgGlnGlu	308
Db	832	GCGAAGGACGCGTCGCGGATGTCGATGATGAGTGAACGCCACCGCGGAGCAGAT	891
Qy	309	LeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaAlaLeuAspAla	328
Db	892	TTCGACGTCCTCGGTGCGGTCAAGGAACCTGGTTCGCGAGCATGCAAGTTCGATCGAC	951
Qy	329	Ser 329	
Db	952	TCG 954	

RESULT 9

ID ADS59454 standard; cDNA; 957 BP.

XX AC ADS59454;

XX DT 02-DEC-2004 (first entry)

XX Bacterial polynucleotide #11441.

DE Recombinant DNA construct; transformed plant; improved plant property;

XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;

KW pathogen tolerance; pest tolerance; plant disease resistance;

KW cell cycle pathway modification; plant growth regulator;

KW homologous recombination; seed oil yield; protein yield; carbohydrate;

KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;

KW bacterial polynucleotide; gene; ss.

XX Bacteria.

OS US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLATER/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

PI WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide

PT for expression of a polynucleotide encoding a polypeptide from a

PT microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 35128; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a

CC promoter functional in a plant cell, where the promoter is positioned to

CC provide for expression of a polynucleotide encoding a polypeptide from a

CC microbial source. The invention also relates to a transformed plant

CC comprising the recombinant DNA construct and a method of producing a

CC transformed plant having an improved property. The plant is a crop plant

CC such as maize or soybean. The method of producing a transformed plant

CC having an improved property comprises transforming a plant with the

CC recombinant DNA construct and growing the transformed plant, where the

CC polynucleotide or polypeptide is useful for improving plant properties.

CC The recombinant DNA construct is useful for producing plants with

CC improved plant properties, e.g. improved cold, heat or drought tolerance,

CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,

CC increased resistance to plant disease, better growth rate by modification

CC of the cell cycle pathway with plant growth regulators, increased rate of

CC homologous recombination, modified seed oil or protein yield and/or

CC content, improved yield by modification of carbohydrate, nitrogen or

CC phosphorus use and/or uptake, by modification of photosynthesis or by

CC providing improved plant growth and development under at least one stress

CC condition. This sequence represents a bacterial polynucleotide used in

CC production. The scope of the invention. Note: The sequence data for this patent did

CC not form part of the printed specification but was obtained in electronic

CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 957 BP; 182 A; 307 C; 273 G; 195 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.74e-84	Length:	957
Score:	873.50	Matches:	167
Percent Similarity:	72.3%	Conservative:	60
Best Local Similarity:	53.2%	Mismatches:	82
Query Match:	51.7%	Indels:	5
DB:	13	Gaps:	1

US-09-390-846-2 (1-330) x ADS59454 (1-957)

QY 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28

DB 7 CGCAAAAAATTGCACCTATTGGTTCTGCGCATGATCGCGCGCACGCTGCGCGCATCTCGCC 66

QY 29 SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMetGly 48

DB 67 AGCTTGAAGGAACATGGCGGATATGCTCTCTTCGACATCGCGGACGCGCATCCCGAGGGC 126

QY 49 LysAlaMetAspIleSerHisAsnSerValValAspThrGlyIleThrValTrpGly 68

DB 127 AAGGCTCTGGATATTGCCAGTCCGGCCGGTGAAGGCTTCAATGCAAGCTCTCCGGC 186

QY 69 SerAsnSerTrpGluCysLeuGlyAlaAspValValIleIleThrAlaGlyIleThr 88

DB 187 GCTTCGCGATTACGCGCGCATCGAAGCGCGACACGCTCTGCATCGTCACCGAGGTGTCGCC 246

QY 89 LysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProValAsnIle 108

DB 247 CGCAAGCCCGGC-----ATGAGCCCGGATGATCTTCTTGGCATCAACCTC 291

QY 109 LysIleMetArgGluValGlyAlaIleLysSerTrpCysProAsnAlaPheValIle 128

DB 292 AAGTCTATGGACAGTTCGCGCGCGCATCAAGAAATATGCTCCGAACGCTTTCGTGATC 351

QY 129 AsnIleThrAsnProLeuAspValMetValAlaLeuGlnGlnSerSerGlyLeuPro 148

DB 352 TGCATCACCAACCGCTCGACGCCATGCTTGGCGGCTGCAGAGTTCTCCGGCGCTCGCG 411

QY 149 HisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMetIle 168

DB 412 AAGAACAAAGTTCGCGCGCATGCTGCGGCTTCGACAGCGACGCTTCCGCTGTTCCTT 471

QY 169 AlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHisGly 188

DB 472 CCCAAGAAATCAACGTTTCGTCAGACGCTCACCGCTTCGTTCTCGCGGTCATGCGC 531

QY 189 AspHisMetValProLeuSerArgTrpAlaThrValAsnGlyIleProLeuSerGluPhe 208

DB 532 GACACCATGTTGCGCGCTCGCCCGTTATTCACCGCTTGGCGGGTTCGCGTTACCGATCTC 591

QY 209 ValLysLysGlyTrpIleLysGlnGluValAspIleValGlnLysThrLysVal 228

DB 592 GTCAAGATGGGCTGTTGACCGCGCAACGCTTTCAGCAGATCATCCAGCGACCGCGTAC 651

QY 229 AlaGlyGlyIleValArgLeuLeuGlyGlnGlySerAlaTrpTrpAlaProGlyAla 248

DB 652 GCGCGCGGGAATTCGTCGCGCTCTTGAAGACCGGCTTCGCGCTATTCGCGCGCGCGCT 711

QY 249 SerAlaIleGlnMetAlaGluSerTrpLeuLysAspArgLysArgValMetValCysSer 268

DB 712 TCGGCTATCGAAATGGCGGATCTTACTCAAGGACAAAGAGCGGTTCTCGCGCTGCT 771

QY 269 CysTrpLeuGlnGlyTrpGlyValGlnAsnHisTrpLeuGlyValProCysValIle 288

DB 772 GCCCACCTTTCCGCCAGTATGGCGTAGACGACATGATGTCGCGCGTTCGCCCATCATC 831

QY 289 GlyClyArgGlyValGluLysIleIleGluLeuLeuThrAlaGlnGluArgGlnGlu 308

DB 832 GGTCCCGCGGATTCGAGCGCGCTCATCGATCGAACTGAACAGGAAGAGAGAGCGCGCC 891

QY 309 LeuGlnGlySerIleAspGluValLysGluMetGlnLysAla 322

DB 892 TTCAGAAATCCGTCGCGCGCTGTCGCTGGTCTTTCGGAAGCC 933

RESULT 10

ADS62502

ID ADS62502 standard; cDNA; 960 BP.

XX

AC ADS62502;

XX

DT 02-DEC-2004 (first entry)

XX

DE Bacterial polynucleotide #14489.

KW Recombinant DNA construct; transformed plant; improved plant property;

XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;

KW pathogen tolerance; pest tolerance; plant disease resistance;

KW cell cycle pathway modification; plant growth regulator;

KW homologous recombination; seed oil yield; protein yield; carbohydrate;

KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;

KW bacterial polynucleotide; gene; ss.

XX OS Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

XX (HINK/) HINKLE G J.

XX (SLAT/) SLATER S C.

XX (CHEN/) CHEN X.

XX (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide

PT for expression of a polynucleotide encoding a polypeptide from a

PT microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 38176; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a

CC promoter functional in a plant cell, where the promoter is positioned to

CC provide for expression of a polynucleotide encoding a polypeptide from a

CC microbial source. The invention also relates to a transformed plant

CC comprising the recombinant DNA construct and a method of producing a

CC transformed plant having an improved property. The plant is a crop plant

CC such as maize or soybean. The method of producing a transformed plant

CC having an improved property comprises transforming a plant with the

CC recombinant DNA construct and growing the transformed plant, where the

CC polynucleotide or polypeptide is useful for improving plant properties.

CC The recombinant DNA construct is useful for producing plants with

CC improved plant properties, e.g. improved cold, heat or drought tolerance,

CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,

CC increased resistance to plant disease, better growth rate by modification

CC of the cell cycle pathway with plant growth regulators, increased rate of

CC homologous recombination, modified seed oil or protein yield and/or

CC content, improved yield by modification of carbohydrate, nitrogen or

CC phosphorus use and/or uptake, by modification of photosynthesis or by

CC providing improved plant growth and development under at least one stress

CC condition, improved lignin production or improved galactomannan

CC production. This sequence represents a bacterial polynucleotide used in

CC the scope of the invention. Note: The sequence data for this patent did

CC not form part of the printed specification but was obtained in electronic

CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 960 BP; 184 A; 307 C; 274 G; 195 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,75e-84	Length:	960
Score:	873.50	Matches:	167
Percent Similarity:	72.3%	Conservative:	60
Best Local Similarity:	53.2%	Mismatches:	82
Query Match:	51.7%	Indels:	5
DB:	13	Gaps:	1

US-09-390-846-2 (1-330) x ADS62502 (1-960)

RESULT 11

ADS62887

ID ADS62887 standard; cDNA; 960 BP.

XX ADS62887;

AC ADS62887;

XX 02-DEC-2004 (first entry)

XX Bacterial polynucleotide #14874.

DE

9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28

CGCAAAAAATTCACCTTATGTTCTGGCATGATCGCGGCAGCTGGGGCATCTCGCC 66

29 SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMetGly 48

AGCCTGAAGGAACATGGCGCATATCGTCTCTTCGACATCGCCGACGCGCATCCGCGAGGC 126

49 LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyGly 68

AGGGTCTGGATATTGCCAGTCCCGCCGGTTGAAGGCTTCAATGCAAGACTCTCCGCC 186

69 SerAsnSerTyrgluCysLeuLysGlyAlaAspValValIleIleThrAlaGlyIleThr 88

GCTTCGATTACCGCCCATCGAAGGCGCAGAGCTGTGCAATCGTCACCGAGGTGTGCC 246

89 LysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProValAsnIle 108

CGCAAGCCCGGC-----ATGAGCCGCGATGATCTTCTTGGCATCAACCTC 291

109 LysIleMetArgGluValGlyAlaAlaIleLysSerTyrgCysProAsnAlaPheValIle 128

AGGTTCATGGAACAGAGTCGCGCGCGCATCAAGAAATATGCTCGAACGCTTTCGTGATC 351

129 AsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeuPro 148

TGATACACCAACCCGCTCGACGCATGCTGTGGCGCTGCAGAGTTCTCGGCGCTGCCG 411

149 HisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMetIle 168

AAGAACAAAGTCTCGTGGCATGGTGTCTCGACAGCGCAGCTTCCGCTTCCGCTTCT 471

169 AlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHisGly 188

GCGAAGAATTCACCGTTCGTCAGGACGTCCCGCTTCTCGCGGTCATGGC 531

189 AspHisMetValProLeuSerArgTyrgAlaThrValAsnGlyIleProLeuSerGluPhe 208

GACACATGTGCGCTCGCCCGTTATTCACCGTTGGCGCGGTTCGCTTACCGATCTC 591

209 ValLysGlyTrpIleLysGlnGluValAspAspIleValGlnLysThrLysVal 228

GTCAAGATGGCTGGTTGACCGCGCAACGCTTGTGAGCAGATCATCCAGCGCACCCGTGAC 651

229 AlaGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrgAlaProGlyAla 248

GGCGCGCGGAAATCGTCGCGCTTTGAAGACCGGCTCGGCTTATTCGCGCGCGCGCT 711

249 SerAlaIleGlnMetAlaGluSerTyrgLeuLysAspArgLysArgValMetValCysSer 268

TCGGCTATCGAATGGCGGAATCTTACTCAAGGACAAGAGCGCGTTCGCCCGCTGCT 771

269 CysTyrgLeuGlnGlyTrpIleGlyValGlnAsnHisTyrgLeuGlyValProCysValIle 288

GCCACCTTTCGGCGCAGTATGGCTAGACGACATGATGTGTCGCGGTGCCCATCATC 831

289 GlyGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGluArgGlnGlu 308

GGTCCGCGGTATCGAGCGCGTTCATCGAGTGAACACTGAACAGGAAGAAGAGCCGCC 891

309 LeuGlnGlySerIleAspGluValLysGluMetGlnLysAla 322

TTCCAGAAATCCGTCGCGCTGTCGCTGTTTTCGGAAGCC 933

XX Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polynucleotide; gene; ss.
 XX
 OS Bacteria.
 XX
 PN US2003233675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAO//) CAO Y.
 PA (HINK//) HINKLE G J.
 PA (SLAT//) SLATER S C.
 PA (CHEN//) CHEN X.
 PA (GOLD//) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 DR
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polynucleotide used in
 CC the scope of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 960 BP; 184 A; 307 C; 274 G; 195 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.75e-84 Length: 960
 Score: 873.50 Matches: 167
 Percent Similarity: 72.3% Conservative: 60
 Best Local Similarity: 53.2% Mismatches: 82
 Query Match: 51.7% Indels: 5
 DB: 13 Gaps: 1
 US-09-390-846-2 (1-330) x ADS62887 (1-960)
 Qy 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28

Db 7 CGCAAAAATATGACCTTATTGGTTCTGGCATGATCGCGGCACGCTCGGCATCTCGCC 66
 Qy SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMetGly 48
 Db 67 AGCTGTAGGAACATGGCGGATATCGTCTCTTCGACATCGCGCGCATCCCGAGGCG 126
 Qy LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyrGly 68
 Db 127 AAGGTCCTGGATATTTGCCAGCTCGCGCGCGTGAAGGCTTCAATGCAAAAGCTCTCGCG 186
 Qy 69 SerAsnSerTyrGluCysLeuLysGlyAlaValValIleIleThrAlaGlyIleThr 88
 Db 187 GCTTCGGATTACGGCGCCATCGAAGCGCACGCTCTGCATCGTCACCGAGGTGTCGCC 246
 Qy 89 LysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProValAsnIle 108
 Db 247 CGCAAGCCCGCG-----ATGAGCCCGGATGATCTTCTTGGCATCAACCTC 291
 Qy 109 LysIleMetArgGluValGlyAlaIleLysSerTyrCysProAsnAlaPheValIle 128
 Db 292 AAGGTCATGGAACACAGTTCGGCGCGCATCAAGAAATATGCTCCGAACGCTTTCGTGATC 351
 Qy 129 AsnIleThrAsnProLeuAspValMetValAlaLeuGlnGluSerSerGlyLeuPro 148
 Db 352 TGCATCACCAACCCCGCTCGACGCCATGCTTGGCGGCTGCAGAAATCTTCGCGGCTCGCG 411
 Qy 149 HisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgArgMetIle 168
 Db 412 AAGAACAGGTTCGCGCATGCTGGCGTTCGACAGCGCAGCTTCGCTGTTCCT 471
 Qy 169 AlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHisGly 188
 Db 472 GCCGAAGAATTCAGGTTTCGCTCGAGCAGCTCACCGCTTCGTTCTCGCGGTTCATGCG 531
 Qy 189 AspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGluPhe 208
 Db 532 GACACCATGGTTCGCGCTCGCGCTTATTCACCGTTCGCGCGCTTCGCTTACCGATCTC 591
 Qy 209 ValLysLysGlyTyrIleLysGlnGluValAspIleValGlnLysThrLysVal 228
 Db 592 GTCAAGTGGCTGGTGGACCGCGCAAGCTCTTGAGCAGATCATCCAGCGCAGCGTGCAC 651
 Qy 229 AlaGlyGlyIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaProGlyAla 248
 Db 652 GGCGGCGGGAATCGTGGCGCTCTTGAAGACCGCTCGGCTTATACGCGCGCGCGCT 711
 Qy 249 SerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCysSer 268
 Db 712 TCGGCTATCGAAATGGCGAATCTTACCTCAAGGACACGAAGCGGCTTCTCGCGCTGCT 771
 Qy 269 CysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysValIle 288
 Db 772 GCCACCTTCGCGCGCGATGTCGCGTACGACATGATGTGTCGGGCTGCCACCATCATC 831
 Qy 289 GlyGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGluArgGlnGlu 308
 Db 832 GGTGCGCGCGGTATCGAGCGCGTTCATCGAGATCGAATCGAAGGACGAAGAGAGAGAGAG 891
 Qy 309 LeuGlnGlySerIleAspGluValLysGluMetGlnLysAla 322
 Db 892 TTCCAGAAATCCGTTCGCGCTGTCGCTGCTTCTTTCGGAAGCC 933
 RESULT 12
 ID ADT45997 standard; cDNA; 933 BP.
 XX
 AC ADT45997;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polynucleotide #20748.
 XX

KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmoesis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polynucleotide; gene; ss.
 XX
 OS Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

XX (HINK/) HINKLE G J.

XX (SLAT/) SLATER S C.

XX (CHEN/) CHEN X.

XX (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide
 XX for expression of a polynucleotide encoding a polypeptide from a
 XX microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 35747; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
 XX promoter functional in a plant cell, where the promoter is positioned to
 XX provide for expression of a polynucleotide encoding a polypeptide from a
 XX microbial source. The invention also relates to a transformed plant
 XX comprising the recombinant DNA construct and a method of producing a
 XX transformed plant having an improved property. The plant is a crop plant
 XX such as maize or soybean. The method of producing a transformed plant
 XX having an improved property comprises transforming a plant with the
 XX recombinant DNA construct and growing the transformed plant, where the
 XX polynucleotide or polypeptide is useful for improving plant properties.
 XX The recombinant DNA construct is useful for producing plants with
 XX improved plant properties, e.g. improved cold, heat or drought tolerance,
 XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 XX increased resistance to plant disease, better growth rate by modification
 XX of the cell cycle pathway with plant growth regulators, increased rate of
 XX homologous recombination, modified seed oil or protein yield and/or
 XX content, improved yield by modification of carbohydrate, nitrogen or
 XX phosphorus use and/or uptake, by modification of photosynthesis or by
 XX providing improved plant growth and development under at least one stress
 XX condition, improved lignin production or improved galactomannan
 XX production. This sequence represents a bacterial polynucleotide used in
 XX the scope of the invention. Note: The sequence data for this patent did
 XX not form part of the printed specification but was obtained in electronic
 XX format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 969 BP; 174 A; 310 C; 309 G; 176 T; 0 U; 0 Other;

XX Alignment Scores:

Pred. No.:	3.11e-83	Length:	969
Score:	862.00	Matches:	174
Percent Similarity:	70.8%	Conservative:	56
Best Local Similarity:	53.5%	Mismatches:	87
Query Match:	51.1%	Indels:	8
DB:	13	Gaps:	2

XX US-09-390-846-2 (1-330) x ADS60073 (1-969)

Qy 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28

Db 7 CGCAACAAGATAGCGCTTATCGGCTCGGGCATGATCGCGCGCACGCTGGCCCATGATC 66

Qy	29	SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMetGly	48
Db	67	GGCCTCAAGACACCTCGCGGACGCTGCTGTGATATGCCGAGGGTATTCGCCAAGGC	126
Qy	49	LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyrgly	68
Db	127	AAGGGGCTCATATCGCGGACGCTGCTCGCGGTCGATGTTCCGATTCGGGCTGACCGGC	186
Qy	69	SerAsnSerTyrglyCysLeuLysGlyAlaAspValValIleIleThrAlaGlyIleThr	88
Db	187	GTCAACGACTATGCGCGCATCGAGGGTGGCGACGCTGTCATGTTACAGCGGGGTGCGG	246
Qy	89	LysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProValAsnIle	108
Db	247	CGCAAGCCTGGC-----ATGAGCGCGACGATCTGCTGGGATCAACCTC	291
Qy	109	LysIleMetArgGluValGlyAlaIleLysSerTyrcysProAsnAlaPheValIle	128
Db	292	AAGGTATGGAAACAGGTGGCGCGGCTGCTGAGAGATATGCGCCCAAGGCTTCGTATC	351
Qy	129	AsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeuPro	148
Db	352	TGCATCACCAATCGCTCGACGCCATGCTGTGGGCGCTGCAGAGATTTCGGGCTGCC	411
Qy	149	HisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMetIle	168
Db	412	AAGACCATGTCGTCGGCATGCGCGGCTGCTGCACAGCGCGCTTCCTGTTTCCTG	471
Qy	169	AlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHisGly	188
Db	472	GCCGAGGAATTCAGGTCGCTCGTCAGGAGCTCACCGCTTCGTGCTCGCGGCGCACGC	531
Qy	189	AspHisMetValProLeuSerArgTyralaThrValAsnGlyIleProLeuSerGluPhe	208
Db	532	GATTCATGTCGCGATGATCGGCTACTCGACGCTATCGGGCATCCGCTGCCCGACCTC	591
Qy	209	ValLysLysGlyTrpIleLysGlnGluValAspIleValGlnLysThrLysVal	228
Db	592	GTCAAGATGGGCTGGACCTCGAAGGAGAGCTCGACCATGCTCCAGCGGCACCGCTGAT	651
Qy	229	AlaGlyGlyGluIleValArgLeuGlyGlnGlySerAlaTyrtyrAlaProGlyAla	248
Db	652	GGCGGGCGCGAGATCGTCGGCTGCTCAAGACCGGCTCGGCTATTACGGCGCGCGGCC	711
Qy	249	SerAlaIleGlnMetAlaGluSerTyrlsLysAspArgLysArgValMetValCysSer	268
Db	712	TCGGCGATTGGCATGGCGGATCTCTACTCAAGGACAAAGACCGCGTGTGCTGCTGCT	771
Qy	269	CysTyrlsLeuGlnGlyTrpIleGlyValGlnAsnHisTyrlsLeuGlyValProCysValIle	288
Db	772	GCCACCTTTTCGGCGCATGTCGCTCAAGGGCACCTATGTCGGCGTCCCGTGTGATC	831
Qy	289	GlyGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGluArgGlnGlu	308
Db	832	GGCGCGCGCGCGTCGAGCGCATCATGAGATCATGAGATCATCAACAGAGCGACACAAAATG	891
Qy	309	LeuGlnGlySerIleAspGluValLysGluMetGlnLysAla-----IleAlaAla	325
Db	892	TTGAGAGATTGGTGGCGACCGTGCAGGGCTGCAGAGGCTGCTGGTCAAGATCGCGCG	951
Qy	326	LeuAspAlaSerLys	330
Db	952	CAGCTCGCTTCGAAG	966
RESULT 14			
ID	ADT42328		
XX	ADT42328 standard; cDNA; 963 BP.		
XX	AC ADT42328;		
XX	DT 02-DEC-2004 (first entry)		

DE Bacterial polynucleotide #17079.

XX Recombinant DNA construct; transformed plant; improved plant property;

KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;

KW pathogen tolerance; pest tolerance; plant disease resistance;

KW cell cycle pathway modification; plant growth regulator;

KW homologous recombination; seed oil yield; protein yield; carbohydrate;

KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;

XX bacterial polynucleotide; gene; ss.

XX OS Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAO//) CAO Y.

XX (HINK//) HINKLE G J.

XX (SLAT//) SLATER S C.

XX (CHEN//) CHEN X.

XX (GOLD//) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide

PT for expression of a polynucleotide encoding a polypeptide from a

PT microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 40766; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a

CC promoter functional in a plant cell, where the promoter is positioned to

CC provide for expression of a polynucleotide encoding a polypeptide from a

CC microbial source. The invention also relates to a transformed plant

CC comprising the recombinant DNA construct and a method of producing a

CC transformed plant having an improved property. The plant is a crop plant

CC such as maize or soybean. The method of producing a transformed plant

CC having an improved property comprises transforming a plant with the

CC recombinant DNA construct and growing the transformed plant, where the

CC polynucleotide or polypeptide is useful for improving plant properties.

CC The recombinant DNA construct is useful for producing plants with

CC improved plant properties, e.g. improved cold, heat or drought tolerance,

CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,

CC increased resistance to plant disease, better growth rate by modification

CC of the cell cycle pathway with plant growth regulators, increased rate of

CC homologous recombination, modified seed oil or protein yield and/or

CC content, improved yield by modification of carbohydrate, nitrogen or

CC phosphorus use and/or uptake, by modification of photosynthesis or by

CC providing improved plant growth and development under at least one stress

CC condition, improved lignin production or improved galactomannan

CC production. This sequence represents a bacterial polynucleotide used in

CC the scope of the invention. Note: The sequence data for this patent did

CC not form part of the printed specification but was obtained in electronic

XX format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 963 BP; 174 A; 333 C; 314 G; 142 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5,41e-82	Length:	963
Score:	850.50	Matches:	164
Percent Similarity:	69.8%	Conservative:	60
Best Local Similarity:	51.1%	Mismatches:	92
Query Match:	50.4%	Indels:	5
DB:	13	Gaps:	1

US-09-390-846-2 (1-330) x ADT42328 (1-963)

9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28

7 CGCGGAGATCGCCCTTATCGCGCGGATGATCGCGCGCACCTGCGCCACATCGCC 66

29 SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMetGly 48

67 GCTCGGAGAGGTGGCGGACGTGATCTCTTCACATCGCGGAGGACCGCGGAGGG 126

49 LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyrGly 68

127 AAGGCCCTCGACATCGCGGAGGCTGCGCGCTGTTCGCGAAGGACGATGGCCCTGAAGGC 186

69 SerAsnSerTyrGluCysLeuLysGlyValAlaAspValValIleIleThrAlaGlyIleThr 88

187 GCCAACGACTACGCCGACATCGCGCGGCTGACGTCTGCATCGTACCGCGGGGTGCG 246

89 LysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProValAsnIle 108

247 CGCAAGCGGGC-----ATGAGCGCGGACGACCTGCTGGGCATCAACCTG 291

109 LysIleMetArgGluValGlyAlaAlaIleIysSerTyrCysProAsnAlaPheValIle 128

292 AAGGTCTATGAAGCGCGTGGCGGAGGATCAAGGCTCACGCCCCCAACGCTTCGTGATC 351

129 AsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeuPro 148

352 TGATACCAACCGCTGGACGGATGTTCTGGGCCCTGCAGCAGTTCTCGGGCTGCG 411

149 HisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMetIle 168

412 AAGGAAAAGGTCTCGGATGCGCGGCTCTCGACTCGCGCGCTTCGCTACTCTCTG 471

169 AlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHisGly 188

472 GCTGAAGCCACCGCGCTCTCGGTGGAGACATCCACGCCCTGGACCCTGGCGGCGCACGGC 531

189 AspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGluPhe 208

532 GACGACATGTGCGATGTGCTGCTACGCGTGGCGCGCTGCGCTGCGGAACTG 591

209 ValLysGlyTrpIleLysGlnGluValAspAspIleValGlnLysThrLysVal 228

592 GTCAAGCAAGGTGGCTGCTCGAGGACAAAGTGGACGCCATCGTCGAGCGCACCGCAAG 651

229 AlaGlyGluLeuValArgLeuGlyGlnGlySerAlaTyrTyrAlaProGlyAla 248

652 GCGCGCGCGAGATCGTGGCCCTCTCAAGACCGGCTCGGCTCTACGCCCGCGCGAG 711

249 SerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCysSer 268

712 AGCGGATCGCATGGCGACCTCGTACTGAGGACAAAGAGCGCGTCTGCGCGTGGCC 771

269 CysTyrLeuGlnGlyIleTyrGlyValGlnAsnHisTyrLeuGlyValProCysValIle 288

772 ACCTACCTGCGCGCCAGTACGGCCTCGAACACCTCTATGTGCGCGTGGCGTGTGATC 831

289 GlyClyArgGlyValGluLysIleIleGluLeuGluThrAlaGlnGluArgGlnGlu 308

832 GCGCGCGCGCGCGGAGAGATCGTGTGAGTTCGAAACCAACGACGACGAGAGAGCGATG 891

309 LeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaLeuAspAla 328

892 TTGCGCAAGTGGTGTGAGTGGTGAAGGCTGTATGAAGGCTGTCAAGGCTCGACAGC 951

329 Ser 329

952 TCG 954

RESULT 15

ADSS6026

ID ADS56026 standard; cDNA; 960 BP.

XX AC ADS56026;

XX 02-DEC-2004 (first entry)
 XX Bacterial polynucleotide #8013.
 XX Recombinant DNA construct; transformed plant; improved plant property;
 XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 XX pathogen tolerance; pest tolerance; plant disease resistance;
 XX cell cycle pathway modification; plant growth regulator;
 XX homologous recombination; seed oil yield; protein yield; carbohydrate;
 XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 XX bacterial polynucleotide; gene; ss.
 XX Bacteria.
 XX US2003233675-A1.
 XX 18-DEC-2003.
 XX 20-FEB-2003; 2003US-00369493.
 XX 21-FEB-2002; 2002US-0360039P.
 XX (CAOY) CAO Y.
 XX (HINK) HINKLE G J.
 XX (SLAT) SLATER S C.
 XX (CHEN) CHEN X.
 XX (GOLD) GOLDMAN B S.
 XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 XX New recombinant DNA construct comprising a promoter positioned to provide
 XX for expression of a polynucleotide encoding a polypeptide from a
 XX microbial source, useful for producing plants with improved properties.
 XX Claim 1; SEQ ID NO 31700; 122pp; English.
 XX The invention relates to a recombinant DNA construct comprising a
 XX promoter functional in a plant cell, where the promoter is positioned to
 XX provide for expression of a polynucleotide encoding a polypeptide from a
 XX microbial source. The invention also relates to a transformed plant
 XX comprising the recombinant DNA construct and a method of producing a
 XX transformed plant having an improved property. The plant is a crop plant
 XX such as maize or soybean. The method of producing a transformed plant
 XX having an improved property comprises transforming a plant with the
 XX recombinant DNA construct and growing the transformed plant, where the
 XX polynucleotide or polypeptide is useful for improving plant properties.
 XX The recombinant DNA construct is useful for producing plants with
 XX improved plant properties, e.g. improved cold, heat or drought tolerance,
 XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 XX increased resistance to plant disease, better growth rate by modification
 XX of the cell cycle pathway with plant growth regulators, increased rate of
 XX homologous recombination, modified seed oil or protein yield and/or
 XX content, improved yield by modification of carbohydrate, nitrogen or
 XX phosphorus use and/or uptake, by modification of photosynthesis or by
 XX providing improved plant growth and development under at least one stress
 XX condition, improved lignin production or improved galactomannan
 XX production. This sequence represents a bacterial polynucleotide used in
 XX the scope of the invention. Note: The sequence data for this patent did
 XX not form part of the printed specification but was obtained in electronic
 XX format from USPTO at seqdata.uspto.gov/sequence.html.
 XX Sequence 960 BP; 169 A; 311 C; 325 G; 155 T; 0 U; 0 Other;
 XX Alignment Scores:
 XX Pred. No.: 8,86e-82 Length: 960
 XX Score: 848.50 Matches: 163
 XX Percent Similarity: 70.7% Conservative: 64
 XX Best Local Similarity: 50.8% Mismatches: 89
 XX Query Match: 50.3% Indels: 5
 XX DB: 13 Gaps: 1

US-09-390-846-2 (1-330) x ADS56026 (1-960)

Qy 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyGlyThrMetAlaPheLeuCys 28
 Db 7 AGACCAAGATTGGCTGCTGCGGGGCGAGATCGGGGCGACGCTCGCCCATCTCGCC 66
 Qy 29 SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMetGly 48
 Db 67 CGCATCAAGGAACATGGGGCGAGCTGCTGTTGCATATCGCGAAGGACCCCTCAGGGC 126
 Qy 49 LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyGly 68
 Db 127 AAGGCTCTGGACATCGCGCAATCGGGCCCTCGGAAGGCTTCGACGGCGTGTGAAGGGC 186
 Qy 69 SerAsnSerTyTrGluCysLeuLysGlyAlaAspValValIleIleThrAlaGlyIleThr 88
 Db 187 GCGAACAGCTACGAGAGATCGCGGGCGCGAGCTCTGCATCGTGACCGGGGGTGGCG 246
 Qy 89 LysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProValAsnIle 108
 Db 247 CGCAAGCCCGGC-----ATGAGCGCGACGACCTGATCGGCATCAACCTC 291
 Qy 109 LysIleMetArgGluValGlyAlaAlaIleLysSerTyCysProAsnAlaPheValIle 128
 Db 292 AAGGTGATGAATCGGTGCGGAGGATCAAGGCCCATCGCGCAACGCTTCGTGATC 351
 Qy 129 AsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeuPro 148
 Db 352 TGCATCACCAACCCGCTCGACCGATGGTCTGGGGCGCTGCAGCAATTCCTCGGGCTCGCG 411
 Qy 149 HisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMetIle 168
 Db 412 GCCGAGAAGTGGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 471
 Qy 169 AlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHisGly 188
 Db 472 TCGGTGAGTTCACACGCTCTCGATCGCGGACGTGACGGCTTCTGTTCTGGGGCGGCATG 531
 Qy 189 AspHisMetValProLeuSerArgTyTrAlaThrValAsnGlyIleProLeuSerGluPhe 208
 Db 532 GACAGATGGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 591
 Qy 209 ValLysLysGlyTrpIleLysGlnGluValAspAspIleValGlnLysThrLysVal 228
 Db 592 GTGCAGATGGCTGGACACCGACGAGAGCTCGACCATCGTGACGCGCACCGCTGAC 651
 Qy 229 AlaGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyTrAlaProGlyAla 248
 Db 652 GCGCGCGCGAGATCGTGGGCGCTGCTGAAGACCGGCTCGGCTTCTACGCGCGCGCAC 711
 Qy 249 SerAlaIleGlnMetAlaGluSerTyTrLeuLysAspArgLysArgValMetValCysSer 268
 Db 712 TCGGCCCATCGAGTGGCGGAAGCCTATCTCAAGACCGAGAGCGTCTGCTGCTGCTGCG 771
 Qy 269 CysTyTrLeuGlnGlyLntTyGlyValGlnAsnHisTyTrLeuGlyValProCysValIle 288
 Db 772 GCCTATGTGACGCGGCTTGGCGCTGAAACGGCATGTATGTGGGGCGTGGCCCATCATC 831
 Qy 289 GlyGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGluArgGlnGlu 308
 Db 832 GCGCGCGCGGAGTCGAGAGATCGTCGACATCAAGCTCAACGACGACGACGACGACGATG 891
 Qy 309 LeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaLeuAspAla 328
 Db 892 TTCGACCAAGTCGTCATCGGTGAAGGGCGCTCGTCGAGGGCTCGCAAGGGGATCGACTCC 951
 Qy 329 Ser 329
 Db 952 TCG 954

Job time : 118 secs



GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 2, 2006, 23:46:13 ; Search time 4597 Seconds
(without alignments)
4080.561 Million cell updates/sec

Title: US-09-390-846-2

Perfect score: 1688

Sequence: 1 MAVFEKTRPKIAVWGSGMI.....GSIDEVKEMOKATAALDASK 330

Scoring table:

	BLOSUM62	Xgapop 10.0	Xgapext 0.5
Xgapop 10.0			0.5
Xgapop 6.0			7.0
Deiop 6.0			7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abs/ABSSWEB spool/US09390846/runat_02032006_104230_4827/app query fasta_1
-DB=GenEmbl -OEMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=spect -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs03p
-USER=US09390846 @CGN 1 1 7415 @runat_02032006_104230_4827 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMOUT=120
-WAE TIMEOUT=30 -THREMS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_cm.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pr.*

9: gb_to.*

10: gb_sts.*

11: gb_ey.*

12: gb_un.*

13: gb_vi.*

14: gb_htg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1688	100.0	1567	2 AY143388	AY143388 Eimeria a
2	1398	82.8	1411	2 AY143390	AY143390 Eimeria m
3	1266	75.0	1876	2 AY143389	AY143389 Eimeria t

4	1072	63.5	2593	2	TGU23207	U33207 Toxoplasma
5	1041.5	61.7	1822	2	TGU35118	U35118 Toxoplasma
6	1034.5	61.3	1785	6	AR279170	Sequence
7	1034.5	61.3	1785	6	AR279171	Sequence
8	905.5	53.6	951	2	AY650028	AY650028 Toxoplasma
9	905.5	53.6	951	2	AY972525	AY972525 Toxoplasma
10	889.5	52.7	1839	1	RLAJ2750	AJ002750 Rhizobium
11	889	52.7	1350	2	AB112429	AB112429 Babesia b
12	881	52.2	951	2	DQ060151	DQ060151 Plasmodium
13	880.5	52.2	31047	1	EX897699	Continuation (20 o
14	876.5	51.9	11111	1	AE009456	AE009456 Brucella
15	873.5	51.7	963	1	AF322647	AF322647 Sinorhizo
16	873.5	51.7	10956	1	AE009211	AE009211 Agrobacte
17	873.5	51.7	11683	1	AE008177	AE008177 Agrobacte
18	873.5	51.7	81384	1	EX897700	Continuation (16 o
19	873.5	51.7	333800	1	SMES91792	AL591792 Sinorhizo
20	868.5	51.5	110000	1	AE017223	Continuation (19 o
21	868	51.4	110000	1	BA000040	Continuation (5 of
22	863.5	51.2	110000	1	AE042918	Continuation (19 o
23	862.5	51.1	349315	1	EX572593	EX572593 Rhodopeu
24	862	51.1	110000	1	BA000012	Continuation (35 o
25	851.5	50.4	963	1	AY181040	AY181040 Methylola
26	851.5	50.4	110000	1	CP000031	Continuation (4 of
27	850.5	50.4	8617	1	AE006023	AE006023 Caulobact
28	847	50.2	898	2	AY486060	AY486060 Plasmodiu
29	838.5	49.7	13627	1	AE008614	AE008614 Rickettsi
30	835	49.5	898	2	AY486059	AY486059 Plasmodiu
31	831.5	49.3	951	2	AB122147	AB122147 Plasmodiu
32	831.5	49.3	951	2	AF251291	AF251291 Plasmodiu
33	831.5	49.3	951	2	PFALDH	M3720 Plasmodium
34	831.5	49.3	110000	14	PFMAL13	Continuation (11 o
35	831.5	49.3	318221	14	PFMAL13P3	AL049184 Plasmodiu
36	829.5	49.1	312430	1	RPX02	AJ23271 Rickettsi
37	826.5	49.0	110000	1	AE017197	Continuation (5 of
38	825.5	48.9	110000	1	CP000053	Continuation (7 of
39	810.5	48.0	220559	1	AE017260	AE017260 Wolbachia
40	808.5	47.9	2210	15	BBU06076	U80676 Botryococcu
41	808.5	47.9	110000	1	AE017321	Continuation (4 of
42	807.5	47.8	951	2	AF323520	AF323520 Plasmodiu
43	807	47.8	898	2	AY486058	AY486058 Plasmodiu
44	806.5	47.8	110000	1	AE017180	Continuation (16 o
45	806.5	47.8	110000	1	AE017180_16	Continuation (17 o

ALIGNMENTS

RESULT 1	AY143388	1567 bp	linear	INV 18-MAY-2004
LOCUS	Eimeria acervulina	lactate dehydrogenase (LDH) mRNA, complete cds.		
DEFINITION	Eimeria acervulina			
ACCESSION	AY143388			
VERSION	AY143388.1	GI:25989636		
KEYWORDS	Eimeria acervulina			
SOURCE	Eimeria acervulina			
ORGANISM	Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;			
REFERENCE	1 (bases 1 to 1567)			
AUTHORS	Schaap, D., Arts, G., Kroeze, J., Niesse, R., Roosmalen-Vos, S.V., Spreeuwenberg, K., Kuiper, C.M., Beek-Verhoeven, N.V.D., Kok, J.J., Kneegtel, R.M.A. and Vermeulen, A.N.			
TITLE	An Eimeria vaccine candidate appears to be lactate dehydrogenase; characterization and comparative analysis			
JOURNAL	Parasitology 128 (6), 603-616 (2004)			
REFERENCE	2 (bases 1 to 1567)			
AUTHORS	Schaap, D.C.			
TITLE	Characterization and cloning of lactate dehydrogenase from three Eimeria species			
JOURNAL	Unpublished			
REFERENCE	3 (bases 1 to 1567)			
AUTHORS	Kok, H.J., van den Boogaart, P., Vermeulen, A.N. and Schaap, D.C.			
TITLE	Direct Submission			
JOURNAL	Submitted (20-AUG-2002) Parasitology R&D, Intervet, Wim de			

Korverstraat, Boxmeer 5830AA, The Netherlands

FEATURES	Location/Qualifiers		
	1..1567		
	/organism="Eimeria acervulina"		
	/mol_type="mRNA"		
	/db_xref="taxon:5801"		
	1..1567		
	/gene="LDH"		
	227..1219		
	/gene="LDH"		
	codon_start=1		
/product="lactate dehydrogenase"			
/protein_id="AA38975.1"			
/db_xref="GI:25989637"			
ORIGIN			
Alignment Scores:			
Pred. No.:	1.76e-133	Length:	1567
Score:	1688.00	Matches:	330.
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0
US-09-390-846-2 (1-330) x AY143388 (1-1567)			
Qy	1	MetAlaValPheGluLysAsnThrArgProLysIleAlaMetValGlySerGlyMetIle	20
Db	227	ATGCGCGTCTTCGAGAAGAAATACACGCCCAAGATTGCTATGTGGGCTCCGGTATGATT	286
Qy	21	GlyGlyThrMetAlaPheLeuCysSerLeuArgGluLeuGlyAspValValLeuPheAsp	40
Db	287	GGAGGCACCATGGCTTCTGTGCGAGCTTGAGGAACTCGAGATGTGTCTCTTCGAC	346
Qy	41	ValValProAsnMetProMetGlyLysAlaMetAspIleSerHisAsnSerSerValVal	60
Db	347	GTGTACCGAATCGCGATGGGGAAGCGATGATATCGCACAAATTCGTCTGGTGGTT	406
Qy	61	AspThrGlyIleThrValTyrGlySerAsnSerTyrGluCysLeuLysGlyAlaAspVal	80
Db	407	GACACGGGTATAACAGTATACGGCTCAAAATTCATACGAGTGTCTTGAAGGGTGGCGACGTA	466
Qy	81	ValIleIleThrAlaGlyIleThrLysIleProGlyLysSerAspLysGluTyrTrpSerArg	100
Db	467	GTAATATAACAGCAGGGATTAACAAAGATACCCGGAAGAGCGATAAAGATGGTCTAGA	526
Qy	101	MetAspLeuLeuProValAsnIleLysIleMetArgGluValGlyAlaAlaIleLysSer	120
Db	527	ATGGATCTATTACCTGTGAATATAAAGATAATAGGAGGTCGTCGACGAATTAATCT	586
Qy	121	TyrCysProAsnAlaPheValIleAsnIleThrAsnProLeuAspValMetValAlaAla	140
Db	587	TACTGTCTTAATGATTGTATTATAATAACAAATCTTTTAGATGTGTAGTGTAGCTGCT	646
Qy	141	LeuGlnGluSerSerGlyLeuProHisIleArgIleCysGlyMetAlaGlyMetLeuAsp	160
Db	647	CTTCAAGAGTCATCAGGACTACCTCATATAGAAATCTGCGGTATGGCTGGGATGCTTGAT	706
Qy	161	SerSerArgPheArgMetIleAlaAspLysLeuGluValSerProArgAspValGln	180
Db	707	AGCTCTCGTTTACACGTATGATAGCTGATAAATTAGAGTCTCTCTCTAGAGATGTACAG	766
Qy	181	GlyMetValIleGlyValHisGlyAspHisMetValProLeuSerArgTyrAlaThrVal	200
Db	767	GGGATGGTCAATAGGTGTACACGGCGATCATATGTGTGCCCCCTAACTAGATATGCAACAGTT	826
Qy	201	AsnGlyIleProLeuSerGluPheValLysLysGlyTrpIleLysGlnGluGluValAsp	220

Db	827	AACGGCATCCCGCTTCTCAGTTTGTAAAGAGGGCTGGATCAAGCAAGAAGATAGAT	886	
Qy	221	AspIleValGlnLysThrLysValAlaGlyGlyGluIleValArgLeuLeuGlyGlnGly	240	
Db	887	GATATCGTTTCAAGAAGCAAGGTCGCTGGAGGAGAGATCGTAGCCCTATTAGCAAGGC	946	
Qy	241	SerAlaTyrTyrAlaProGlyAlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAsp	260	
Db	947	TCGTCTTACTATGCTCCAGGGCTTCAGCTATTTCAGATGGCTGAGAGCTATCTTAAAGGAT	1006	
Qy	261	ArgLysArgValMetValCysSerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHis	280	
Db	1007	AGAAAGAGAGTGTGTTTCTCTTCTACTTTCAGAGGACCAATATGCTGTACAGATCAC	1066	
Qy	281	TyrLeuGlyValProCysValIleGlyGlyArgGlyValGluLysIleIleGluLeuGlu	300	
Db	1067	TACTTAGGAGTACCTTGTGTATCGGTGGGAGAGGTGTTGAGAAGATATTATTGAGTTAGAA	1126	
Qy	301	LeuThrAlaGlnGluArgGlnGluLeuGlnGlySerIleAspGluValLysGluMetGln	320	
Db	1127	TTGACCGCACAAAGAAAGACAGGAGCTTCAGGGATCTATCGATGAGGTTAAGGAGATGCAG	1186	
Qy	321	LysAlaIleAlaAlaLeuAspAlaSerLys	330	
Db	1187	AAGGCTATTGCTGCTCTTCATGATCATCAAG	1216	
RESULT 2				
LOCUS	AY143390	1411 bp	mRNA	linear INV 18-MAY-2004
DEFINITION	Eimeria maxima lactate dehydrogenase (LDH) mRNA, complete cds.			
ACCESSION	AY143390			
VERSION	AY143390.1	GI:25989640		
KEYWORDS				
SOURCE	Eimeria maxima			
ORGANISM	Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae; Eimeria.			
REFERENCE	1 (bases 1 to 1411)			
AUTHORS	Schaap,D., Arts,G., Kroeze,J., Niessen,R., Roosmalen-Vos,S.V., Spreuwenberg,K., Kuiper,C.M., Beek-Verhoeven,N.V.D., Kok,J.J., Knegtel,R.M.A. and Vermeulen,A.N.			
TITLE	An Eimeria vaccine candidate appears to be lactate dehydrogenase; characterization and comparative analysis			
JOURNAL	Parasitology 128 (6), 603-616 (2004)			
REFERENCE	2 (bases 1 to 1411)			
AUTHORS	Schaap,D.C.			
TITLE	Characterization and cloning of lactate dehydrogenase from three Eimeria species			
JOURNAL	Unpublished			
REFERENCE	3 (bases 1 to 1411)			
AUTHORS	Niessen,R. and Schaap,D.C.			
TITLE	Direct Submission			
JOURNAL	Submitted (20-AUG-2002) Parasitology R&D, Intervet, Wim de Korverstraat, Boxmeer 5830AA, The Netherlands			
FEATURES	Location/Qualifiers			
source	1..1411			
	/organism="Eimeria maxima"			
	/mol_type="mRNA"			
	/db_xref="taxon:5804"			
gene	1..1411			
	/gene="LDH"			
CDS	1..1411			
	/gene="LDH"			
	/codon_start=1			
	/product="lactate dehydrogenase"			
	/protein_id="AA38977.1"			
	/db_xref="GI:25989641"			

INNIYKTKQAGGDIVRLLGQGSAYVAPGTSAILMAESYLLKDKRLFVSSCYLNGQYN

ORIGIN
VNNHVLGVPCIIIGKIEIIEIDLNQEEKLLQSSIDEVLEMQATAALDAGR"

Alignment Scores:
Pred. No.: 5.8e-109 Length: 1411
Score: 1398.00 Matches: 266
Percent Similarity: 89.4% Conservative: 29
Best Local Similarity: 80.6% Mismatches: 35
Query Match: 82.8% Indels: 0
DB: 2 Gaps: 0

US-09-390-846-2 (1-330) x AY143389 (1-1411)

Qy 1 MetAlaValPheGluLysAsnThrArgProLysIleAlaMetValGlySerGlyMetIle 20
Db 1 ATGCGGTTTGTAGCAAAATTAAGCGCCTTAAGATTCCTCGTGGCTCGCATGATT 60

Qy 21 GlyGlyThrMetAlaPheLeuCysSerLeuArgGluLeuGlyAspValValLeuPheAsp 40
Db 61 GGAGGCACCATGGCATCTTATGACGCTTAAAGAAATAGGAGATGTTTCTCTTTGAC 120

Qy 41 ValValProAsnMetProMetGlyLysAlaMetAspIleSerHisAsnSerSerValVal 60
Db 121 GTTGCTACCTAATATGCTATGGGAAAGCGATGATTTATGTCATATTCATCTGTAGTA 180

Qy 61 AspThrGlyIleThrValTyrGlySerAsnSerTyrGluCysLeuLysGlyAlaAspVal 80
Db 181 GACAACGGTATACTGTATACGGATCTAATAGCTATGAATGCTTAACCTAACCGCAGATGA 240

Qy 81 ValIleIleThrAlaGlyIleThrLysIleProGlyLysSerAspLysGluTyrSerArg 100
Db 241 GTAATAATTAAGTCTGAGGTATACAAAATAACCCGTAAGTGAATAAGATGTCACAGA 300

Qy 101 MetAspLeuLeuProValAsnIleLysIleMetArgGluValGlyAlaAlaIleLysSer 120
Db 301 ATGGATTATTACAGTAATATTAATAATATGCTGAGGTAGGAGGAGCAATAAAGAA 360

Qy 121 TyrCysProAsnAlaPheValIleAsnIleThrAsnProLeuAspValMetValAlaAla 140
Db 361 TATTGTCTTAATGCTAATTAATAATAATAACAATCCTTTAGATGTAATGGTTGCTGCT 420

Qy 141 LeuGlnGluSerSerGlyLeuProHisHsIleArgIleCysGlyMetAlaGlyMetLeuAsp 160
Db 421 GTACAAGAGCAAGCAATATACCTAAACATATATGCTGTCATGCGGAGGATGCTAGAT 480

Qy 161 SerSerArgPheArgMetIleAlaAspLysLeuGluValSerProArgAspValGln 180
Db 481 TCATCAAGACTAAGAGAAATGTTGCTGATTTGTTGATGTATCTCTCTCATGACGTACAG 540

Qy 181 GlyMetValIleGlyValHisGlyAspHisMetValProLeuSerArgTyrAlaThrVal 200
Db 541 GGTATGTAATAGGTGTACATGAGAGATAATATGTTACTCTTATGCGCTTATATAACTATT 600

Qy 201 AsnGlyIleProLeuSerGluPheValLysGlyTyrIleLysGlnGluValAsp 220
Db 601 AATGGTATACCTATACCAAGATTTTATTAATAAGGATTAATAATAAGGAGAAATAAAT 660

Qy 221 AspIleValGlnLysThrLysValAlaGlyGlyGluIleValArgLeuLeuGlyGlnGly 240
Db 661 AATATATATATAGACTAAGCAAGAGGAGGATATAGTAATATTATTAGGCAAGGA 720

Qy 241 SerAlaTyrTyrAlaProGlyAlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAsp 260
Db 721 AGTGCATATTATGACCAAGGTTACTAGTGAATATTAATGACAGAAAGTACCTTAAGAT 780

Qy 261 ArgLysArgValMetValCysSerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHis 280
Db 781 AAAAAAGATATTATTTCTTCTTCTTGTACCTTAATGACAAATAATATTAATAATCAT 840

Qy 281 TyrLeuGlyValProCysValIleGlyGlyArgGlyValGluLysIleIleGluLeuGlu 300
Db 841 TAICTAGGTGCTCTTGTATTAATAGGAGAAAGGATAGAACAAATCATAGAAATTAGAT 900

Qy 301 LeuThrAlaGlnGluArgGlnGluLeuGlnGlySerIleAspGluValLysGluMetGln 320
Db 901 CTTAATCAGGAGAGAAAAGTTACTACAAGGATCTATTGATGAGGTACTAGATGCAA 960

Qy 321 LysAlaIleAlaAlaLeuAspAlaSerLys 330
Db 961 AAGCTATTGCTGCTCTTGTATGAGGAAG 990

RESULT 3
AY143389

LOCUS Eimeria tenella lactate dehydrogenase (LDH) mRNA linear INV 18-MAY-2004
ACCESSION AY143389
VERSION AY143389.1 GI:25989638
KEYWORDS
SOURCE Eimeria tenella
ORGANISM Eimeria tenella
Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;
Eimeria.
REFERENCE 1 (bases 1 to 1876)
AUTHORS Schaap,D., Arts,G., Kroeze,J., Niessen,R., Roosmalen-Vos,S.V.,
Spreuwerberg,K., Kuiper,C.M., Beek-Verhoeven,N.V.D., Kok,J.J.,
Knegtel,R.M.A. and Vermeulen,A.N.
TITLE An Eimeria vaccine candidate appears to be lactate dehydrogenase;
characterization and comparative analysis
JOURNAL Parasitology 128 (6), 603-616 (2004)
REFERENCE 2 (bases 1 to 1876)
AUTHORS Schaap,D.C.
TITLE Characterization and cloning of lactate dehydrogenase from three
Eimeria species
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1876)
AUTHORS Arts,G., Kroeze,H. and Schaap,D.C.
TITLE Direct Submission
JOURNAL Submitted (20-AUG-2002) Parasitology R&D, Intervet, Wim de
Korverstraat, Boxmeer 5830AA, The Netherlands

FEATURES
Location/Qualifiers
1..1876
/organism="Eimeria tenella"
/mol_type="mRNA"
/db_xref="taxon:5802"
1..1876
/gene="LDH"
175..1170
/gene="LDH"
/codon_start=1
/product="lactate dehydrogenase"
/protein_id="AA38976.1"
/db_xref="GI:25989638"
/translation="MAVFEKRRPKIALVSGMIGTGMFLCSLRELGDVVLFDVVPN
MPAGKALDCHTAADVADNGVRVQAGNSYASLEGADVVIITAGITKAAGKSDQMSRKD
LLPVNKKILREVCAATKQCPHAFVININPLDMVAALREAAAGLPAARVCGWAGVLD
SARPKLLADRLGVSPEDVQAMVLGVHGDNNWPLSPATVNGVPLGELARQCHISEAE
IREVERQTRAAGDIVRLUGQSSAIPAGAAVAAMAEYLDKQKRVFCVSCYLEGGPG
VRHGCLGVPCVWAGGVVERVIELPLDAREALQLQASIDVREMHRLQAADAABAE"

ORIGIN

Alignment Scores:
Pred. No.: 1.27e-97 Length: 1876
Score: 1266.00 Matches: 235
Percent Similarity: 86.9% Conservative: 51
Best Local Similarity: 71.4% Mismatches: 43
Query Match: 75.0% Indels: 0
DB: 2 Gaps: 0

US-09-390-846-2 (1-330) x AY143389 (1-1876)

Qy 1 MetAlaValPheGluLysAsnThrArgProLysIleAlaMetValGlySerGlyMetIle 20
Db 175 ATGCGGTTTTCAGAAAGGTCCGCGCGCCCAAGATCGCGCTCGTGGCTCCGCGATGATC 234

Qy 21 GlyGlyThrMetAlaPheLeuCysSerLeuArgGluLeuGlyAspValValLeuPheAsp 40

```
Db 235 GGAGGCACGATGGGTTTCCTGTGCAGCGCTCGGAGCTCGGAGACGCTGCTGCTTCGAC 294
Qy 41 ValValProAsnMetProMetGlyLysAlaMetAspIleSerHisAsnSerSerValVal 60
Db 295 GTCGTGCCGAACATGCCCGCGGGAAGGCGCTGCACCTGTGCCACACGCGGAGTGCGG 354
Qy 61 AspThrGlyIleThrValTyrgLysSerAsnSerTyrgLysLeuLysGlyAlaAspVal 80
Db 355 GACACGCGCTGAGGTTGCAGGCGCGACGACGCTGCAGGCTGAGGCGCGGACGCTG 414
Qy 81 ValIleIleThrAlaGlyIleThrLysIleProGlyLysSerAspLysGluTrpSerArg 100
Db 415 GTGATATAACGCGCAGCAATAACGAGCGCGCGGGAAGGACCAAGAGTGGTCCGCG 474
Qy 101 MetAspLeuLeuProValAsnIleLysIleMetArgGluValGlyAlaAlaIleLysSer 120
Db 475 AAAGACCTGCTGCGGCTGAGGCTGAGGCGCGAAGTCTCGGGAAGTCCGCGCCATCAAGCAG 534
Qy 121 TyrCysProAsnAlaPheValIleAsnIleThrAsnProLeuAspValMetValAlaAla 140
Db 535 TTCTGCCCGCAGCGCTTCGTCATCAATCACCACCCCTCGACGTGATGTCGCGCG 594
Qy 141 LeuGlnGluSerSerGlyLeuProHisHisArgIleCysGlyMetAlaGlyMetLeuAsp 160
Db 595 CTCGCGAAGCGCGGACTCCCGCAGCCGCGCTCTGCGGATGCGCGCGTCTCTCGAC 654
Qy 161 SerSerArgPheArgArgMetIleAlaAspLysLeuGluValSerProArgAspValGln 180
Db 655 TCCGCGCGCTTCCGACAGCTCTCGCGACCGCTCGCGCTCAGCCCCCGCGACGCTCCAG 714
Qy 181 GlyMetValIleGlyValHisGlyAspHisMetValProLeuSerArgTyrgAlaThrVal 200
Db 715 GCATGTGCTCTCGCGCTCCACGCGCACAATGTCGCGCTGAGTCGCTTCGCGACGGTG 774
Qy 201 AsnGlyIleProLeuSerGluPheValLysGlyTyrgIleLysGlnGluValAsp 220
Db 775 AACGGCGTCCGCTCGCGAGCTCGCGCGCCGCGGTGGATCTCGGAGCGGAGATCCGC 834
Qy 221 AspIleValGlnLysThrLysValAlaGlyGlyGluIleValArgLeuGlyGlnGly 240
Db 835 GAGGTGAGCGCCAGACGCGCGCGCGCGGCGACATTTGTCGCTGCTGCGGCGAGGC 894
Qy 241 SerAlaTyrgAlaProGlyAlaSerAlaIleAlaIleMetAlaGluSerTyrgLeuLysAsp 260
Db 895 TCTGCGTACTTCCGCGCGCGCTGCGCGCTGCCATGCCGAGGCTTACTTGAAGAC 954
Qy 261 ArgLysArgValMetValCysSerCysTyrgLeuGlnGlyGlnTyrgIleValGlnAsnHis 280
Db 955 CAGAAAGCGCTCTTCGTGTGTTCTGCTACTTGGAGGGGCTTACGCGCGTGGCGCAC 1014
Qy 281 TyrLeuGlyValProCysValIleGlyCysArgGlyValGluLysIleIleGluLeuGlu 300
Db 1015 TGTGTGGCGTGCCTTCCGTGCTGCGGCGCGCGCTTACGCGCGTCACTGAGCTGCCG 1074
Qy 301 LeuThrAlaGlnGluArgGlnGluLeuGlnGlySerIleAspGluValLysGluMetGln 320
Db 1075 CTCAGCCCCCGAGGCGCGAGCTTTCAGGCTTCCATTGACGAGGTCCCGAGATGCAC 1134
Qy 321 LysAlaIleAlaAlaLeuAspAlaSer 329
Db 1135 CGCCAGCTCGCTGCTGCCGAGCTGCT 1161

RESULT 4
TGU23207
LOCUS
DEFINITION
Toxoplasma gondii lactate dehydrogenase mRNA, complete cds.
ACCESSION
U23207
VERSION
U23207.1
GI:975280
KEYWORDS
Toxoplasma gondii
Toxoplasma gondii
Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
Sarcocystidae; Toxoplasma.
```

```
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. -2593
/organism="Toxoplasma gondii"
/mol_type="mRNA"
/strain="ME49"
/db_xref="taxon:5811"
/clone="LDH2"
/dev_stage="bradyzoite"
/lab_host="Mus musculus"
1. .268
5'UTR
CDS
269. .1249
/EC_number="1.1.1.27"
/codon_start=1
/product="lactate dehydrogenase"
/protein_id="AAC48663.1"
/db_xref="GI:975281"
/translation="MTGTVSRKKIAMIGSMIGTMGYLCVLRLEADVLFVFDVVTGM
PEGKALDDSQATSIADTNVSVTSANQYKIASGVVITAGLTQVPGSKDKMSRNDL
LFPNAKIREVAQVKYCPAFVIVTNPLDMVKCFHEAGSLPKNMVCGMANVLDS
ASERFIADOLEISPRDQATVGTGHDHMLPLARVTVTSGPLREFIKKGKWEAKL
ASIVERTKAGGEIVRLGQGSAYYAPALSALTMAQAFKDKERVLPCSVYCGEGYGL
HDMFGLPFAVGGGIEVLELTHEBQCFKSVDDVVVLELKSIALUG"
1250. .2593
3'UTR
ORIGIN
Alignment Scores:
Pred. No.: 5,26e-81 Length: 2593
Score: 1072.00 Matches: 199
Percent Similarity: 80.2% Conservative: 56
Best Local Similarity: 62.6% Mismatches: 63
Query Match: 63.5% Indels: 0
DB: 2 Gaps: 0
US-09-390-846-2 (1-330) x TGU23207 (1-2593)
Qy 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyGlyThrMetAlaPheLeuCys 28
Db 290 AGAAAAAAGATTGCAATGATTGTTCTGCGATGATAGGAGGAACCATGGCTACCTTTGT 349
Qy 29 SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMetGly 48
Db 350 GTGCTTCGGGAGCTAGCCGACGCTCGTACTATTTCGACGTTGTAAACAGGCATGCCAGAAGG 409
Qy 49 LysAlaMetAspIleSerHisAsnSerValValAspValValIleIleThrAlaGlyIleThr 68
Db 410 AAGCGCTTGGATGATTCAAGCGCAACAGCATTTGTCACACGACACGCTGAGCGTGACGAC 469
Qy 69 SerAsnSerTyrgLysCysLysGlyAlaAspValValIleIleThrAlaGlyIleThr 88
Db 470 GCNAACCATGATGAGAGATCGCCGATCGGATGCTAATAATACTACGAGGCTGACC 529
Qy 89 LysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProValAsnIle 108
Db 530 AAGGTACCGGGGAAGAGTGCACAGGAGTGGAGCAGAACGACCTTTTACCGTTCAATGCA 589
Qy 109 LysIleMetArgGluValGlyAlaAlaIleLysSerTyrgCysProAsnAlaPheValIle 128
Db 590 AAAATCATTCGAGAGGTAGCGGAGGAGTGAAGAGTACTGCCCCGCTTCTTTTGTAAT 649
Qy 129 AsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeuPro 148
```

```

Db      650 GTAGTGACAAACCGCTTGAAGTGGTAAAGTGGTCTTTCATGAAGCGGCGGTACCA 709
Qy      149 HishisargilleCysGlyMetAlaGlyMetLeuAspSerArgPheArgMetIle 168
Db      710 AAAAAACATGGTTTCGGAATCGCAATGTGTAGATTGGCAGCATTTAGACGCTTCATT 769
Qy      169 AlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHisGly 188
Db      770 CGGATCAGCTGGAGATATCACTCGAGATATTCAAGCTACTGTATTCGGCAGCACACGGT 829
Qy      189 AspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGluPhe 208
Db      830 GACCACATGCTGCTCTCGCGGTTATGTGACGGTGAGTGGTTCCTCCACTCGGAGATTTC 889
Qy      209 vallylsysGlyTyrIleLysGlnGluValAspAspIleValGlnLysThrLysVal 228
Db      890 ATTAATAAAGAAATACCTGAGCGCAAACTGGCAGAGATCGTAGAGCGGTACAAAGAAA 949
Qy      229 AlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaProGlyAla 248
Db      950 GCTGGGGGGAGATTGTGAGACTTTTGGGGCAAGGGTCAGCGTACTATGCTCCAGCCCTT 1009
Qy      249 SerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCysSer 268
Db      1010 TCTGCTATACAAATGGCCAGGCAATCTTAAGACGAGAGAGAGTCTCCCGTGCAGT 1069
Qy      269 CysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysValIle 288
Db      1070 GTGTATTGCCAGGTGAGTATGGCTTCACGATATGTTATTGTTGGCGGCGAGTTATA 1129
Qy      289 GlyGlyArgGlyValGluLysIleGluLeuGluIleThrAlaGlnGluArgGlnGlu 308
Db      1130 GGCGGTGGGGAATCGACGAGTTCATCGAATCGAGCTGCACATGAGGAACAAGAGTGT 1189
Qy      309 LeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaLeu 326
Db      1190 TTCAGGAAGTCAGTTGATGAGTGGTGGATGAATTGAATGAGTTTAGCGGCGCTG 1243

RESULT 5
LOCUS   TGU35118 1822 bp mRNA linear INV 05-FEB-1997
DEFINITION Toxoplasma gondii lactate dehydrogenase mRNA, complete cds.
ACCESSION U35118
VERSION   U35118.1 GI:1695771
KEYWORDS
SOURCE    Toxoplasma gondii
          Toxoplasma gondii
          Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
          Sarcocystidae; Toxoplasma.
REFERENCE 1 (bases 1 to 1822)
AUTHORS   Yang, S. and Parmley, S.F.
TITLE     Toxoplasma gondii expresses two distinct lactate dehydrogenase
          homologous genes during its life cycle in intermediate hosts
JOURNAL   Gene 184 (1), 1-12 (1997)
PUBMED   9016946
REFERENCE 2 (bases 1 to 1822)
AUTHORS   Parmley, S.F.
TITLE     Direct Submission
JOURNAL   Submitted (30-AUG-1995) Stephen F. Parmley, Immunology and
          Infectious Diseases, Research Institute, Palo Alto Medical
          Foundation, 860 Bryant Street, Palo Alto, CA 94301, USA
FEATURES
         Location/Qualifiers
           1..1822
             /organism="Toxoplasma gondii"
             /mol_type="mRNA"
             /strain="ME49"
             /db_xref="taxon:5811"
             /clone="LDH1"
             /dev_stage="tachyzoite"
             /lab_host="Mus musculus"
           1..143
             /EC_number="1.1.1.27"

```

5'UTR
CDS

```

/codon_start=1
/product="lactate dehydrogenase"
/protein_id="AAC47443.1"
/db_xref="GI:1695772"
/translation="MAPALVQRKKVKAMIGSMIGTMGYLCALRELADVVLVYDVVKG
MPGKALDLSHVTSVVDNTVSVRAEYSYEAALTGADCVIVTAGLTKVPGKPDSEWSRN
DLAPFNKIIIREIGQALIKYCPKTFIIVTNPLDCVKKVCEASGVPTNMLCGNACML
DSGRFRYVADALSVSPDQATVITGHGDCMFLVRIITVNGPIQKFIKIDGVVTEK
GLBEIAHETKVSGGEIVRFLGQSGAYAPAAVSAMATSFINDKRRVIPCISVYCNSEY
QLKDMPTIGLPAVIGGAGIERVIELENEEEKQFQKSVDDVMALNKAAVALQA"
1134..1822
3'UTR
ORIGIN
Alignment Scores:
Pred. No.: 1,3e-78 Length: 1822
Score: 1041.50 Matches: 194
Percent Similarity: 77.6% Conservative: 55
Best Local Similarity: 60.4% Mismatches: 71
Query Match: 61.7% Indels: 1
DB: 2 Gaps: 1
US-09-390-846-2 (1-330) x TGU35118 (1-1822)
Qy      9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
Db      168 AGAAAGAGGTGGCCATGATTGGCTCTGGCATGATTGGTCGCACTATGGGTACCTGTGC 227
Qy      29 SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMetGly 48
Db      228 GCTCTCGTGAGCTCGCTGACGCTCTCTAGATGTTGTCAAGGTATGCCGAGGCT 287
Qy      49 LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyrGly 68
Db      288 AAGGCTCTTACGACCTGAGCCATGACCTCGTGGTGCACACCAACGTTTCGTCGGTGC 347
Qy      69 SerAsnSerTyrGlu---CysLeuLysGlyAlaAspValValIleThrAlaGlyIle 87
Db      348 GAGTACTCTTACGAGGCGCGCTCACCGGTGCGGACTGCTGTATTCGTACCGCGGCTG 407
Qy      88 ThrValIleProGlyLysSerAspLysGluTyrSerArgMetAspLeuLeuProValAsn 107
Db      408 ACCAAGGTGCGGGCAAGCCGAGCTCGAGTGAGCGGAGAACGATCTGCTCCGCTTCAAC 467
Qy      108 IleLysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAlaPheVal 127
Db      468 TCGAGATCATTCGCGAGATCGGTCAACAATCAAGAGTACTGCCCAAGAGCTTCATC 527
Qy      128 IleAsnIleThrAsnProLeuAspValMetValAlaLeuGlnGluSerSerGlyLeu 147
Db      528 ATCGTGTCAACCAACCGCTGGACTGCATGGTCAAGGTTATGTGCGAGGCTCTTGGCGTC 587
Qy      148 ProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMet 167
Db      588 CCACCAACATGATCTCGGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 647
Qy      168 IleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHis 187
Db      648 GTCCGCGAGCGGCTGTCTGTCTCTCCCGCAGCTCCAGGCCACCGTATCGGCACACAC 707
Qy      188 GlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGlu 207
Db      708 GCGGACTGTCATGCTCCCGCTTGTCCGGTACATTACCGTGAACGGCTACCGCATCCAGAAG 767
Qy      208 PheValLysLysGlyTyrIleLysGlnGluValAspAspIleValGlnLysThrLys 227
Db      768 TTCATCAAGGACGCGGTAGTCACGAGAGACAGCTCGAGGAGATCGTGCAGCACACAAA 827
Qy      228 ValAlaGlyGlyLurIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaProGly 247
Db      828 GTGCTGGCGGCGAGATCGTCCGCTTCCTCGGCGAGGTTCCGCTTACTACGCCCGCCGCC 887
Qy      248 AlaSerAlaIleGlnMetAlaGluSerTyrTyrLeuLysAspArgLysArgValMetValCys 267

```

```
Db 888 GCATCCGCTGCGCCATGGCAACATCTCTTGAACGACGAAAGCGCGTCATCCCGTGC 947
Qy 268 SerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysVal 287
Db 948 AGTGTGCTGCTCAACGAGGAGTACGGCTTGAAGACATGTTTCATGTGCTCTCCCGCGCTC 1007
Qy 288 IleGlyGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGluArgGln 307
Db 1008 ATTGAGCGCCGCGCATCGAGCGCGTCTATCGAGCTCGAGCTGAACGAGGAGGAGAAG 1067
Qy 308 GluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaLeuAsp 327
Db 1068 CAGTTCAGAAAGTCCGTCGACGAGTCTATCGCGCTCAACAGCGGTTGCTGCTCTTCAG 1127
Qy 328 Ala 328
Db 1128 GCG 1130

RESULT 6
LOCUS AR279170 1785 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 311 from patent US 6514694.
ACCESSION AR279170
VERSION AR279170.1 GI:29713813
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1785)
AUTHORS Milhausen,M.J.
TITLE Methods for the detection of encysted parasites
JOURNAL Patent: US 6514694-A 311 04-FEB-2003;
Heska Corporation; Fort Collins, CO
FEATURES
    source
        1. 1785
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.: 4.95e-78 Length: 1785
Score: 1034.50 Matches: 193
Percent Similarity: 77.3% Conservative: 55
Best Local Similarity: 60.1% Mismatches: 72
Query Match: 61.3% Indels: 1
Gaps: 6
DB: 1

US-09-390-846-2 (1-330) x AR279170 (1-1785)
Qy 9 ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
Db 109 AGAAGAAGGTGGCCATGATTTGGCTCTGGCATGATTGGTGGCACTATGGGCTACCTGTGC 168
Qy 29 SerLeuArgGluLeuGlyAspValValPheAspValValProAsnMetProMetGly 48
Db 169 GCTCTCGTGCAGTCGCTGACGCTGCTCTACGATGTTCTCAAGTATGCCCGAGGT 228
Qy 49 LyAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyrGly 68
Db 229 AAGGCTTTCAGCTGAGCCATGTGACCTCGGTGTCACCAACAGTTTCGGTCCGTGCT 288
Qy 69 SerAsnSerTyrGlu---CysLeuLysGlyAlaAspValValIleIleThrAlaGlyIle 87
Db 289 GAGTACTCTTACGAGGCGCGCTCACCAGTGGCGACTGCGTTATCTACCGCGGTCTG 348
Qy 88 ThrLysIleProGlyLysSerAspLysGluLysTrpSerArgMetAspLeuLeuProValAsn 107
Db 349 ACCAAGTGGCGGCAAGCCGACTCCGAGTGGAGCGGAACGATCTGCTCCGCTTCAAC 408
Qy 108 IleLysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAlaPheVal 127
Db 409 TCGAAGATCATTCGCGAGATCGGTGAGAACATCAAGAAGTACTGCCCCAGACCTTCATC 468
```

```
Qy 128 IleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeu 147
Db 469 ATCGTGTGCACCAACCCGCTGACATGATGGTCAAGGTCAATGTCGAGGCCTCTGGCGTC 528
Qy 148 ProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMet 167
Db 529 CCGACCAACATGATCTGCGGTATGGCTGTCATGCTGACTCTGGTCTGCTCCGCCGATAC 588
Qy 168 IleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHis 187
Db 589 GTCGCCGAGCGCGTGTCTCTCTCCCGCAGCTCCAGGCCACCGTCAATCGGCACACAC 648
Qy 188 GlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGlu 207
Db 649 GCGGACTGCTGCTCCGCTTGTCCGGTACATTACCGTGAACGACTACCGATCCAGAAG 708
Qy 208 PheValLysGlyTrpIleLysGlnGluValAspAspIleValGlnLysThrLys 227
Db 709 TTTCATCAGGACGGCGTAGTACGCGAAGAGAGCTCGAGGAGATCGCTGAGCAGACCAA 768
Qy 228 ValAlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaProGly 247
Db 769 GTGCTCGCGCGGAGATCTCGCTTCTCCGCGCAGGTTCCGCTTACTACGCCCCGCC 828
Qy 248 AlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCys 267
Db 829 GCATCCGCTGTCGCCATGGCAACATCTTCTTGAACGACGAAAGCGCTCATCCCGTGC 888
Qy 268 SerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysVal 287
Db 889 AGTGTGCTACTGCAACGAGAGTAGCGCTTGAAGGACATGTTTCATGCTCTCCGCCCGTC 948
Qy 288 IleGlyGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGluArgGln 307
Db 949 ATTGGAGCGCGCGGATCGAGCGCTCATCGAGCTCGAGCTCAACAGGAGGAGAAGAAG 1008
Qy 308 GluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaLeuAsp 327
Db 1009 CAGTTCAGAAAGTCCGTCGACGAGCTCATGGCGCTCAACAGGCGGTTCGCTCTTCAG 1068
Qy 328 Ala 328
Db 1069 GCG 1071

RESULT 7
LOCUS AR279171
DEFINITION Sequence 313 from patent US 6514694.
ACCESSION AR279171
VERSION AR279171.1 GI:29713814
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1785)
AUTHORS Milhausen,M.J.
TITLE Methods for the detection of encysted parasites
JOURNAL Patent: US 6514694-A 313 04-FEB-2003;
Heska Corporation; Fort Collins, CO
FEATURES
    source
        1. 1785
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.: 4.95e-78 Length: 1785
Score: 1034.50 Matches: 193
Percent Similarity: 77.3% Conservative: 55
Best Local Similarity: 60.1% Mismatches: 72
Query Match: 61.3% Indels: 1
Gaps: 6
DB: 1
```

US-09-390-846-2 (1-330) x AR279171 (1-1785)		Toxoplasma gondii strain RH malate dehydrogenase mRNA, complete cds.	
Qy	9	ArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeu	Cys 28
Db	109	AGAAGAAGTGGGCATGTTGGTCTGGCATGTTGGTGGCACTATGGGTACCTGTGC	168
Qy	29	SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMetGly	48
Db	169	GCTCTCGTGAGCTCGTGACGTGTTCTTACGATGTTGCAAGGTATCCCGAGGGT	228
Qy	49	LysAlaMetAspIleSerHisAsnSerValValAspThrGlyIleThrValTyrGly	68
Db	229	AAGGCTCTGACCTGAGCATGTGACCTCGTGTGCGACACCAAGTTTCGTCGCT	288
Qy	69	SerAsnSerTyrGlu---CysLeuLysGlyAlaAspValValIleThrAlaGlyIle	87
Db	289	GAGTACTCTTACGAGGCGCGCTCACCGTGGGACTCGCTTATCGTTACCGCGGTG	348
Qy	88	ThrLysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProValAsn	107
Db	349	ACCAGGTGCGGCAAGCCGATCCGAGTGGAGCGGAACGATCTGCTCCCGTTCAAC	408
Qy	108	IleLysIleMetArgGluValGlyAlaIleLysSerTyrCysProAsnAlaPheVal	127
Db	409	TCGAAGATCATTCGGAGATCGGTGAGACATCAAGTACTGCTCCCAAGACCTTCATC	468
Qy	128	IleAsnIleThrAsnProLeuAspValMetValAlaLeuGlnGlnSerSerGlyLeu	147
Db	469	ATCGTGTGTACCAACCCGCTGACATGTCATGTCGATGTCGAGGCGCTCTGGCGTC	528
Qy	148	ProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMet	167
Db	529	CCGACCACATGATCTCGCGTATCGCTGTCATGTCGACTCTGCTCGCTTCCCGCGATC	588
Qy	168	IleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHis	187
Db	589	GTCCCGCAGCGGTGTCTGTCTCTCCCGGACGTCAGGCCACCGTCATCGGCACAC	648
Qy	188	GlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGlu	207
Db	649	GGCGACTGCATGTCGCGCTTGTCCGGTACATTACCGTGAACGACTACCGCATCAGAAG	708
Qy	208	PheValLysLysGlyTyrIleLysGlnGluValAspAspIleValGlnLysThrLys	227
Db	709	TTTCATCAAGGACGCGTAGTCACGAGAAGCAGCTCGAGGAGATCGCTGAGCACACAA	768
Qy	228	ValAlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaProGly	247
Db	769	GTGTCTGGCGGAGATCGCTCGCTTCTCGCGCAGGGTTCGCTTACTAGCCCCCGCC	828
Qy	248	AlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCys	267
Db	829	GCATCCGCTGTGCGCATGCGCAACATCTCTTTGAACGACGAAAGCGGTATCCCGTGC	888
Qy	268	SerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysVal	287
Db	889	AGTGTGTACTGCAACGGAGAGTACGGCTTGAAGACATGTTTCATTGTCTCCCGCGCTC	948
Qy	288	IleGlyGlyArgGlyValGluLysIleIleGluLeuLeuThrAlaGlnGluArgGln	307
Db	949	ATTGGAGCGCCGATCGAGCGGTATCGAGCTCGAGCTGAAACGAGGAGGAGAAAG	1008
Qy	308	GluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaLeuAsp	327
Db	1009	CAGTTCAGAGTCCGTGCGAGCTATGCGCGTCAACAGGCGGTGTGCTCTTCAG	1068
Qy	328	Ala 328	
Db	1069	GCG 1071	
RESULT 8		951 bp mRNA linear INV 07-JUL-2004	
AY650028			
LOCUS			

DEFINITION

Toxoplasma gondii strain RH malate dehydrogenase mRNA, complete cds.

ACCESSION

AY650028

VERSION

AY650028.1 GI:49617504

KEYWORDS

SOURCE

ORGANISM

Toxoplasma gondii

REFERENCE

1 (bases 1 to 951)

AUTHORS

Shen, C., Zhan, X., He, A. and Li, Z.

TITLE

Direct Submission

JOURNAL

Submitted (10-JUN-2004) Department of Parasitology, Zhongshan Medical College, Sun Yat-Sen University, Guangzhou, Guangdong 510089, China

FEATURES

Location/Qualifiers

source

1..951

CDS

/organism="Toxoplasma gondii"

product

/product="malate dehydrogenase"

protein_id

/protein_id="AA67462.1"

db_xref

/db_xref="GI:49617505"

translation

/translation="MSRRKIGLGGNIGATLALLSAVKELGVDVMDVVDLPQKGC

VGAIKQYCPNAFVICITNP

VGAIKQYCPNAFVICITNPDMVYLREKCGLPKHKVCGMAGVLSARLRTFLSR

LNVSDIHALVMGHDNVP

LNVSDIHALVMGHDNVPFRFTTGGIIPLELVGMISQQEVDVIVQTRNG

GGISVLLTKTSAPFAA

GGISVLLTKTSAPFAAAGVLMAYLAKDRKRVLPAAALNGEYGVGDMYGVPCV

IGAGGVEKIVELDTPE

IGAGGVEKIVELDTPEKKMFERSVESVTKLLAAAPKSA"

Alignment Scores:

Pred. No.: 1,91e-67 Length: 951

Score:

905.50 Matches: 173

Percent Similarity:

73.0% Conservative: 57

Best Local Similarity:

54.9% Mismatches: 80

Query Match:

53.6% Indels: 5

DB:

2 Gaps: 1

US-09-390-846-2 (1-330) x AY650028 (1-951)

Qy

8 ThrArgProLysIleAlaMetValGlySerGlyMetIleGlyGlyThrMetAlaPheLeu 27

Db

4 TCTCGAAGAAAATCGCCCTCATCGCGCGGCGCAACATCGCGCCACGCTTGCCTTCTC 63

Qy

28 CysSerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMet 47

Db

64 TCCGCTGTCAAGGAACATCGCGGACGTCGTTCATGTTCCGACGTCGTCAGGACCTCCCGCAA 123

Qy

48 GlyLysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyr 67

Db

124 GGAATAATCGCTCGATCTACCATCGGATGCGGATGCGGATGCGGATGCGGATGCGGATG 183

Qy

68 GlySerAsnSerTyrGluCysLeuLysGlyAlaAspValValIleThrAlaGlyIle 87

Db

184 GGCTCGAACCACTACGACGTCCTCAAGGATGCGGATGCGGATGCGGATGCGGATGCGGATG 243

Qy

88 ThrLysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuProValAsn 107

Db

244 CCTCGGAAGCCAGGT-----ATGCTCGCGACGACCTGCTGGCGATTAAAC 288

Qy

108 IleLysIleMetArgGluValGlyAlaIleLysSerTyrCysProAsnAlaPheVal 127

Db

289 GCGAAAATCATGGCCAAAGTTGGAAGCCATCAAGCAGTACTGCCCCAACGCAATTCGTC 348

Qy

128 IleAsnIleThrAsnProLeuAspValMetValAlaLeuGlnGlnSerSerGlyLeu 147

Db

349 ATTGTGATCATCAAGATCCATCGATGTGATGCTTACATCTCCGGAATAAATCGCGCTTG 408

QY	148	ProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMet	167
Db	409	CTTCCCAACAAGTTTGGCGCATGGCGGGCTCTCGACTCAGCTCGGCTTCGCACGCTTT	468
QY	168	IleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHis	187
Db	469	CTCTCGAGCGCTCAACGCTCTCCGTCGATGACATCCAGCCCTCGTCATGGGTGGCCAT	528
QY	188	GlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGlu	207
Db	529	GGAGACACCATGGTGGCCCTCCCTCGATTCCACCACTGTGGGAGGCATCCCCCTGCCTGAG	588
QY	208	PheValLysLysGlyTyrIleLysGlnGluValAspIleValGlnLysThrLys	227
Db	589	CTGGTGAAGATGGGTATGATTTCTCAACAAGAGTCGACGATATCGTCCAACGCACTCGC	648
QY	228	ValAlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaProGly	247
Db	649	AATGGAGGTGGAGAAATCGTCTCATTCGTAAGACGGGCTCTGCTTCTTCGCTCCCGCG	708
QY	248	AlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCys	267
Db	709	GCTCGGGCGCTTGTGATGGCCGAGGCGGTACTGAAGAGCCGAAACGGGTGCTGCCTTGT	768
QY	268	SerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysVal	287
Db	769	CGCGCTATTTTGAACGGAGAGTACGGTGTCAAGGACATGATGTGGCGGTGCCGTGCGTG	828
QY	288	IleGlyGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGluArgGln	307
Db	829	ATCGCGCGGGCGGCGTCGAGAAGATTGTGCAATTGACCTTGACGCTTGAGGAGAAAG	888
QY	308	GluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAla	322
Db	889	ATGTTTCAGCGCTCGGTTGAAAGTGTGAAGACGCTTCTCGCGGCT	933
RESULT	9		
LOCUS	AY972525	951 bp DNA linear	INV 16-APR-2005
DEFINITION	Toxoplasma gondii malate dehydrogenase gene, complete cds.		
ACCESSION	AY972525		
VERSION	AY972525.1 GI:62465590		
KEYWORDS	Toxoplasma gondii		
SOURCE	Toxoplasma gondii		
ORGANISM	Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Sarcocystidae; Toxoplasma.		
REFERENCE	1 (bases 1 to 951)		
AUTHORS	Chan, M. and Sim, T. S.		
TITLE	Functional characterization of an alternative [lactate dehydrogenase-like] malate dehydrogenase in Plasmodium falciparum		
JOURNAL	Parasitol. Res. 92 (1), 43-47 (2004)		
PUBMED	14598170		
REFERENCE	2 (bases 1 to 951)		
AUTHORS	Shen, C., Zhan, X., Zheng, X., Zheng, B., He, A., Li, Z. and Zheng, H.		
TITLE	Cloning and characterization of the malate dehydrogenase gene from Toxoplasma gondii		
JOURNAL	Zhongguo Ren Shou Gong Huan Bing Za Zhi 21 (2005) In press		
REFERENCE	3 (bases 1 to 951)		
AUTHORS	Shen, C., Zhan, X., Zheng, X., Zheng, B., He, A., Li, Z. and Zheng, H.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-MAR-2005) Department of Parasitology, Zhongshan Medical College, Sun Yat-sen University, Guangzhou, Guangdong 510089, China		
FEATURES	Location/Qualifiers		
source	1..951		
	/organism="Toxoplasma gondii"		
	/mol_type="genomic DNA"		
	/strain="RH"		
	/db_xref="taxon:5811"		
	/dev_stage="tachyzoite"		
mRNA	<1..>951		
	/product="malate dehydrogenase"		

CDS	1..951		
/codon_start=1			
/product="malate dehydrogenase"			
/protein_id="AA83290.1"			
/db_xref="GI:62465591"			
/translation="MSRRKIGLIGGGNIGATLALLSAVKELGDVVMFVVDLPQSKC			
LDLYQUTPISGVDFRFGNSNDYSVLKDDADVIITAGVPRKPGMSRDDLLAIKAKINGQ			
VGEAIKQYCPNFAVICITNPLDMVYILREKGLPHKVCVMAGVLDLSARLRTFLSER			
LNVSDVDLHALVMGGHDTMPLPRFTTVGGIPLPLVMGMISQOEVDDIVQRTNRG			
GGEIVSLKLTGSAPFAAAGVLMAYAYLKDRKRVLPCAAAYLNGEIVGKDMVGVPCV			
IGAGGVEKIVELDLTPPEKMFERSVESVTKLLAAAPKSA"			
ORIGIN			
Alignment Scores:			
Pred. No.: 1	91e-67	Length: 951	
Score: 905.50	Matches: 173		
Percent Similarity: 73.0%	Conservative: 57		
Best Local Similarity: 54.9%	Mismatches: 80		
Query Match: 53.6%	Indels: 5		
DB: 2	Gaps: 1		
US-09-390-846-2 (1-330) x AY972525 (1-951)			
QY	8	ThrArgProLysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeu	27
Db	4	TCTCGAAGAAAAATCGGCCTCATCGCGCGCGCAACATCGCGCCACGCTTGCCTTCTC	63
QY	28	CysSerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMet	47
Db	64	TCGGTGTCAAGAACTCGGCGAGCTGTCATGTTTCGACGCTGTCAGGACCTCCCGCA	123
QY	48	GlyLysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyr	67
Db	124	GGAAAAATGCTCGATCTGTACCGATTGATCTCGATCTCTGGAGTTGACGTTGCTTCGAG	183
QY	68	GlySerAsnSerTyrGluCysLeuLysGlyAlaAspValValIleIleThrAlaGlyIle	87
Db	184	GGCTCGAAGCACTACAGCGTCTCTCAAGGATCGGATGTCATCATCATCGTCACTGCGAGGGGTG	243
QY	88	ThrLysIleProGlyLysSerAspLysGluTyrSerArgMetAspLeuProValAsn	107
Db	244	CCTCGAAGCCAGGT-----ATGCTCGCGACGACCTGCTGGCGATTAAAC	288
QY	108	IleLysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAlaPheVal	127
Db	289	GGAAAAATCATGGGCCAAGTTGGAGAACCATCAAGCAGTACTGCCCAACGCACTTCGTC	348
QY	128	IleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeu	147
Db	349	ATTTCGATCACGAATCCACTCGATGTGATGGTCTACATCTCTCCGCGAAAAATGCGGCTTG	408
QY	148	ProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMet	167
Db	409	CCTCCCAACAAAGTTTGGCGCATGGCGCGCTCTCGACTCAGCTCGGCTTCGACGTTT	468
QY	168	IleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHis	187
Db	469	CTCTCGAGCGTCTCAACGCTCTCCGTCGATGACATCCAGCCCTCGTCATGGGTGGCCAT	528
QY	188	GlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGlu	207
Db	529	GGAGACACCATGGTGGCCCTCCCTCGATTCCACCACTGTGGGAGGCATCCCCCTGCCTGAG	588
QY	208	PheValLysLysGlyTyrIleLysGlnGluValAspIleValGlnLysThrLys	227
Db	589	CTGGTGAAGATGGGTATGATTTCTCAACAAGAGTCGACGATATCGTCCAACGCACTCGC	648
QY	228	ValAlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaProGly	247
Db	649	AATGGAGGTGGAGAAATCGTCTCATTCGTAAGACGGGCTCTGCTTCTTCGCTCCCGCG	708
QY	248	AlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCys	267


```

Db      709  GCTGGGCGCTTTGATGGCGAGCGGTACCTGAAGGACCGCAAGCGGTGCTGCTTGT 768
Qy      268  SerCysTyrLeuGlnGlyClnTyrGlyValGlnHnHstYrLeuGlyValProCysVal 287
Db      769  GCCCCTATTTTGAACGGAGTACGGTGTCAAGGACATGTATGTGGCGTCCCGTGGT 828
Qy      288  IlegGlyATgGlyValGluYalLeileGluLeuGluLeuThrAlaGlnGluArgGln 307
Db      829  ATCGGCGGCGCGGTGCGAAGATTGCGAATTGGACTTGACCTTGAGCGCTTCTCGCGCT 888
Qy      308  GluLeuGlnGlySerIleAspGluValYsGluMetGlnIlyAla 322
Db      889  ATGTTGAGCGCTCGTGTGAAGTGTGAAGCGCTTCTCGCGCT 933

RESULT 10
RLAJ2750
LOCUS
DEFINITION
  Rhizobium leguminosarum mdh gene and partial succ gene and partial
  gene encoding putative ATPase.
ACCESSION
  AJ002750
VERSION
  AJ002750.1 GI:2624393
KEYWORDS
  ATPase; malate dehydrogenase; mdh gene; succ gene; succinyl-CoA
  synthetase beta-subunit.
SOURCE
  Rhizobium leguminosarum
  Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
  Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
  1
  Walshaw,D.L., Wilkinson,A., Mundy,M., Smith,M. and Poole,P.S.
  Regulation of the TCA cycle and the general amino acid permease by
  overflow metabolism in Rhizobium leguminosarum
  Microbiology (Reading, Engl.) 143 (Pt 7), 2209-2221 (1997)
  9245810
  2 (bases 1 to 1839)
  Poole,P.S., Alloway,D. and Smith,M.
  Regulation of the mdh-suc operon of Rhizobium leguminosarum
  Unpublished
  3 (bases 1 to 1839)
  Direct Submission
  Submitted (12-NOV-1997) Poole P.S., Division of Microbiology,
  University of Reading, School of AMS University of Reading,
  Whiteknights, Reading, RG6 6AJ, UK
  Location/Qualifiers
    1..1839
    /organism="Rhizobium leguminosarum"
    /mol_type="genomic DNA"
    /strain="384"
    /sub_species="viciae"
    /db_xref="taxon:384"
    <1..288
    /codon_start=1
    /transl_table=1
    /product="putative ATPase"
    /protein_id="CAA05716.1"
    /db_xref="GI:2624394"
    /db_xref="InterPro:IPR005654"
    /db_xref="UniProt/TREMBL:O33524"
    /translation="VDFLAIAERFDTVFVDHILPLGPKERNOIKRFLILVDITFYDHAV
    RLXISAAMPEELLQRRGTGEGFEDTASRLPMSRAEYALHHEKRAAE"
    466..1428
    /gene="mdh"
    466..1428
    /gene="mdh"
    /codon_start=1
    /transl_table=1
    /product="malate dehydrogenase"
    /protein_id="CAA05717.1"
    /db_xref="GI:2624395"
    /db_xref="GOA:O33525"
    /db_xref="InterPro:IPR001236"
    /db_xref="InterPro:IPR001252"
    /db_xref="InterPro:IPR001557"

```

```

/db_xref="UniProt/Swiss-Prot:O33525"
/translation="MARKKIALICSGMIGGTLAHLAGLKEGLDILVDFDIADIGIPQGGK
LDISQSPVEGFDVNLTGASDYSAIEAGDVCITVAGVARKPMGRDLDLLGNLKMEO
VGAGIKKYAPNAFVICITNPLDAMWALQKFSGLPANKVYGMAGVLDSRSRFLFLAKE
FNYSVQDVTAFTVLGGHDTWPLARYSTVGIGIPLDILVTWGTWTKERLEETIORTRDG
GABIGLLTKTSAYYAPAAIAEMAESYLKDKKRVLPAAHLSSQYGVKDMVYGVPTV
IGAGVERIIEIDLNTEKEAFDKSCAVAGLCEACINIAPAK"
1465..>1839
/gene="succ"
1465..>1839
/gene="sucC"
/codon_start=1
/transl_table=1
/product="succinyl-CoA synthetase beta sub-unit"
/protein_id="CAA05718.1"
/db_xref="GI:2624396"
/db_xref="InterPro:IPR011761"
/db_xref="UniProt/TREMBL:O33526"
/translation="MNIHEYQAKALLKGYGAPVAEGVALIKVEAEAAAKSLPGLPYV
VRSQIHAGGRGKFKFELSPEAKGVLAKSIDEVVAHAKEMLGNTLVTTAQTGEAXKQ
VNRLYTEDGADIARELYCSLLVD"

ORIGIN
Alignment Scores:
Pred. No.:      9,96e-66      Length:      1839
Score:          889.50      Matches:      172
Percent Similarity: 72.9%      Conservative: 59
Best Local Similarity: 54.3%      Mismatches: 81
Query Match:    52.7%      Indels:      5
DB:             1          Gaps:      1

US-09-390-846-2 (1-330) x RLAJ2750 (1-1839)

Qy      6    LysAsnThrArgProIlyIleAlaMetValGlySerGlyMetIleGlyIleThrMetAla 25
Db      463  AAAATGCGCGGTAAACAGATCGCACTATTGTTCTGGCATGATTGTTGGCACGCTGGCG 522
Qy      26  PheLeuCysSerLeuArgGluLeuGlyAspValValLeuPheAspValProAsnMet 45
Db      523  CACCTCGCGCGCTGAAGAACTGGCGCATGCTCTCTTCGACATCGCGACGCGCAT 582
Qy      46  ProMetGlyIlyAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThr 65
Db      583  CCCACGGGCAAGGTTCTGATATTTCCAGTCGTCGCGCGTTCGAAGCGTTTCGACGCAAT 642
Qy      66  ValTyrGlySerAsnSerTyrGluCysLeuLysGlyAlaAspValValIleIleThrAla 85
Db      643  CTGACGGCGCCAGCGACTATTTCGGCATCGAAGCGCTGACGTTCTCATCTCGTCACGCG 702
Qy      86  GlyIleThrIlySileProGlyIlySerAspLysGluTrpSerArgMetAspLeuLeuPro 105
Db      703  GGCCTCGCGCGCAAGCGCGC-----ATGAGCGCGATGACCTTCTCGGC 747
Qy      106  ValAsnIleIlySileMetArgGluValGlyAlaAlaIleIlySerTyrCysProAsnAla 125
Db      748  ATCAACCTCAAGTTCATGGAGCAGGTGCGCGCGCATCAAGAAGTATGCGCGCAACGCT 807
Qy      126  PheValIleAsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSer 145
Db      808  TTGATGATCGCATCAACATCCGTCGACCCATGCTTCGGCGCTTCGCAAGATTTTTC 867
Qy      146  GlyLeuProHisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArg 165
Db      868  GGTCTTCGCGCAACAAGGTGTCGCGCATGCGCGCGCTTCTCGACTCTCGCTTCGCT 927
Qy      166  ArgMetIleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGly 185
Db      928  CTTTTCCTCGCCCAAGGAATTCACACGTGTCGCGTACAGACGTCACGCGCTTCGTTCTCG 987
Qy      186  ValHisGlyAspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeu 205
Db      988  GGCCACGCGCACACGATGGTGGCGCGCTGCGCCCTACTCTGACGCGTTCGCGCGCTC 1047

```

Qy 206 SerGluPheValLysLysGlyTyrPheLysGlnGluValAspAspIleValGlnLys 225
 Db 1048 ACCGATCTGTCACCATGGCTGGGTACCAGAGGCGCTCGAAGAGATCATCCAGGC 1107
 Qy 226 ThrLysValAlaGlyGlyGluValArgLeuLeuGlyGlnGlySerAlaTyrAla 245
 Db 1108 ACCGTCAGCGCGCGCGGAANTCGTCGGCTGCTGAAGACCGCTCGGCTATTACGCC 1167
 Qy 246 ProGlyAlaSerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMet 265
 Db 1168 CCGCCCGCTCGCGATCGAGATGCGGAATCTTACCTCAAGGAGCAAGAGCGGTTCTG 1227
 Qy 266 ValCysSerCysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValPro 285
 Db 1228 CCCTGGCGCGCCACCTCTCGGTCAGTACGCGCTCAAGGACATGATATGTCGGGTTCCC 1287
 Qy 286 CysValIleGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGlu 305
 Db 1288 ACCGTCATCGCGCGCGCGGTGAGCGCATCATCGATCGATCTCAACAGACCGAG 1347
 Qy 306 ArgGlnGluLeuGlnGlySerIleAspGluValLysGluMetGlnLysAla 322
 Db 1348 AAGGAAGCCTTCGACAAGTCGTCGCGCGCTGCTCGGAGCC 1398
 RESULT 11
 LOCUS AB112429 1350 bp mRNA linear INV 08-JUL-2004
 DEFINITION Babesia bovis mRNA for L-lactate dehydrogenase, complete cds.
 ACCESSION AB112429
 VERSION AB112429.1 GI:31790621
 KEYWORDS
 SOURCE Babesia bovis
 ORGANISM Babesia bovis
 Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Babesiidae;
 Babesia.
 REFERENCES
 1 Bork.S., Okamura,M., Boonchit,S., Hirata,H., Yokoyama,N. and Igarashi,I.
 Identification of Babesia bovis L-lactate dehydrogenase as a potential chemotherapeutic target against bovine babesiosis Mol. Biochem. Parasitol. 136, 165-172 (2004)
 2 (bases 1 to 1350)
 YOKOYAMA,N. and IGARASHI,I.
 Direct Submission
 Submitted (13-JUN-2003) Noaki Yokoyama, National Research Center for Protozoan Diseases, Obihiro University of Agriculture and Veterinary Medicine, Inada-cho, Obihiro, Hokkaido 080-8555, Japan (E-mail:yokoyama@obihiro.ac.jp, Tel:81-155-49-5649, Fax:81-155-49-5643)
 FEATURES
 Location/Qualifiers
 1..1350
 /organism="Babesia bovis"
 /mol_type="mRNA"
 /db_xref="taxon:5865"
 gene 1..1350
 /gene="BbLDH"
 CDS 50..1042
 /gene="BbLDH"
 /codon_start=1
 /product="L-lactate dehydrogenase"
 /protein_id="BAC77691.1"
 /db_xref="GI:31790622"
 /translation="MYKNILSTWAPVKRNKISLIGSGNIGGWVAYLAQLKELGDVVL
 FDTAPLGAALDMDHNAIYDTSQVIGTYSYEDTAGSDVCIITAGLAKLPKNSDD
 EWSRDDIYVNSKMTIGENIKKYAPNAFVICITNPLDMVKMLKSTGFPKNKVVG
 MGLLDSRRMCHYIADLRVNPYVHGSCIGHGDSIMPLTNHVTNGIPIQRFIERG
 EITQAEUDKIAERTIGSGMELVOLYNGNSAFFAPATAIEMASAYLSDKESIVCSY
 LESEYGHNDVYLGTPAIIAGNGIEKIITLKLSAEQAKLDASVKETRRLEALVG"
 ORIGIN
 Alignment Scores: 7.44e-66 Length: 1350
 Pred. No.: 889.00 Matches: 166
 Score:

Percent Similarity: 72.4% Conservative: 62
 Best Local Similarity: 52.7% Mismatches: 87
 Query Match: 52.7% Indels: 0
 DB: 2 Gaps: 0
 US-09-390-846-2 (1-330) x AB112429 (1-1350)
 Qy 9 ArgProIysIleAlaMetValGlySerGlyMetIleGlyThrMetAlaPheLeuCys 28
 Db 92 AGGAAACAAATACGCTTATTGGCTCCGGAACATCGGTGGTGTATGGCTTACCTTGCA 151
 Qy 29 SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMetGly 48
 Db 152 CAACTGAGGAACCTTGGCGATGTAGTACTCTTCGACATTGCACCAAAATTTGGGTGCTGCC 211
 Qy 49 LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyrGly 68
 Db 212 AAGGCTTTGGATATCATGCAGCAATGCAATGCACTTACGATACCTCCCAAAATGTTATTGGT 271
 Qy 69 SerAsnSerTyrGluCysLeuLysGlyAlaAspValValIleIleThrAlaGlyIleThr 88
 Db 272 ACAACTAGCTTACGAAGATATCGCCGTAGCGAGCTTTGCATCATCTGCTGGTTTGCA 331
 Qy 89 LysIleProGlyLysSerAspLysGluTyrPheArgMetAspLeuLeuProValAsnIle 108
 Db 332 AAACCTTCCCAACAAGACGATGATGATGGAGCGGTGATGTTGGTGGCCCAACAGC 391
 Qy 109 LysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAlaPheValIle 128
 Db 392 AAGATTATGTTCACTATCGTGAAAAATATCAAGAAGTACGCTCCCAACGCTTTGTGATA 451
 Qy 129 AsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeuPro 148
 Db 452 TGCACTACATAACCATTTGGACGTGTCAGATGCTCAAGATGCTCAATCCACCGATTTCCC 511
 Qy 149 HisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgArgMetIle 168
 Db 512 AAGAACAAGGTGTTGGTATGGTGGACTTTGGATTCTTCCAGATGTGCCACTACATT 571
 Qy 169 AlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHisGly 188
 Db 572 GCCGATAAACTCAGGGTCAACCTTAGGTATGTACACGGATCTCTGTCATTTGGTGGCCAGGT 631
 Qy 189 AspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGluPhe 208
 Db 632 GACAGCATGATTCCTGACTAACCACTGCTCAAGGTATCCCATCCCAAGGTTTC 691
 Qy 209 ValLysLysGlyTyrIleLysGlnGluGluValAlaAspIleValGlnLysThrLysVal 228
 Db 692 ATTGAAGGGAGAAATTACCCAGGCGCACTTTGACAAGATTGCTGACAGCTACCATGGA 751
 Qy 229 AlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaProGlyAla 248
 Db 752 TCTGTATGGAACCTTCTGAGCTTTACGGTAACGGATCTGCTTCTTCGCGCCAGCTACC 811
 Qy 249 SerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCysSer 268
 Db 812 GCTGCCATTGAATGGCATCTGCTTACCTACGTGATAAGAGGTCTGTAAATGTCTGCTCA 871
 Qy 269 CysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysValIle 288
 Db 872 TGCTACTTGGAGGTGAATATGACACACAGATGATATCTTGGCACTCTCTGCCATTATT 931
 Qy 289 GlyGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGluArgGlnGlu 308
 Db 932 GCGCCCAACGGAATGAAAAAATCATCACCTTTAAACTAGTCTGAAGCAAGCAAG 991
 Qy 309 LeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIle 323
 Db 992 TTGGATGCTTCGTCAGGAATATCAGGAGGTTCGAGGCTTTGGTT 1036
 RESULT 12
 DQ060151

Alignment Scores:	2.28e-65	Length:	951
Pred. No.:	891.00	Matches:	176
Score:		Conservative:	48
Percent Similarity:	71.1%	Mismatches:	89
Best Local Similarity:	55.9%	Indels:	2
Query Match:	52.2%	Gaps:	1
DB:	2		

US-09-390-846-2 (1-330) x BX897699_19 (1-31047)
Continuation (20 of 20) of BX897699 from base 1900001 (BX897699 Bartonella henselae str)

Alignment Scores:		1900001 1931047	
Pred. No.:	2.03e-63	Length:	31047
Score:	880.50	Matches:	169
Percent Similarity:	73.2%	Conservative:	66
Best Local Similarity:	52.6%	Mismatches:	81
Query Match:	52.2%	Indels:	5
DB:	1	Gaps:	1
US-09-390-846-2 (1-330) x BX897699_19 (1-31047)			
Qy	6	LysAsnThrArgProLysIleAlaMetValGlySerGlyMetIleGlyGlyThrMetAla 25	
Db	15425	AAAAATGCGACGAAAGAAATAGCTCTTATGGTTCGGGTATGATTGGGGTACTTTTAGCA 15366	
Qy	26	PheLeuCysSerLeuArgGluLeuGlyAspValValLeuPheAspValProAsnMet 45	
Db	15365	CATATATTGCATTAAAGAACTTGGTGATGCTCTTATTTGATATGCAAGGTATTA 15306	
Qy	46	ProMetGlyLysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThr 65	
Db	15305	CCACAAGTAAAGCTCTTGATATCGCGCAATCTTCACCAAGTTGACGGTTCATGATGCAGT 15246	
Qy	66	ValTyrglySerAsnSerTyrglyCysLeuLysGlyAlaAspValValIleIleThrAla 85	
Db	15245	TTAACCGGTGTAATCTCTATGATGATGATGAAGTGCCGATGTTGTTATTTGTAACAGCA 15186	
Qy	86	GlyIleThrLysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuPro 105	
Db	15185	GGTGTTCGAAAGAACCCGGT-----ATGAGCCGTGATGACCTTTTGGGC 15141	
Qy	106	ValAsnIleLysIleMetArgValGlyAlaIleLysSerTyrglyCysProAsnAla 125	
Db	15140	ATTAATTTAAAGGTAATGGAACAAGTTGGTGGGGATTAATAAATATGCACTTCACGA 15081	
Qy	126	PheValIleAsnIleThrAsnProLeuAspValMetValAlaLeuLeuGlnGluSerSer 145	
Db	15080	TTCTGTTATTTGTTATACAAATCTCTTGATGCAATGCTGGGCTTTGCAAAAATTTTCA 15021	
Qy	146	GlyLeuProHisArgIleCysGlyMetAlaGlyMetLeuAspSerArgPheArg 165	
Db	15020	GGTCTTCCACACAAAAGTGGTGGTATGCTGCTGATTCCTCGATTGACACGGTTTCGT 14961	
Qy	166	ArgMetIleAlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGly 185	
Db	14960	CATTCTTATCAGAGGAATTTAAGATCTCTGTTAAAGATGTGACTGCATTCGTTTAGGA 14901	
Qy	186	ValHisGlyAspHisMetValProLeuSerArgTyrglyAlaThrValAsnGlyIleProLeu 205	
Db	14900	GGACACGGTGATTCATGCTGCTTGGTGGTATTTCACAGTTGGTGGTATTTCCTA 14841	
Qy	206	SerGluPheValLysLysGlyTyrIleLysGlnGluValAspAspIleValGlnLys 225	
Db	14840	CCGATCTTGTAATAATGGCGTGGACCGACGCAAGAAAGATTGATCAAAATTAATCAACGT 14781	
Qy	226	ThrLysValAlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrglyAla 245	
Db	14780	ACCGTGATGGTGGTCAGAAATTTGTTAGTTGTTAAACACAGGTTCTGCTTTTATGCA 14721	
Qy	246	ProGlyAlaSerAlaIleGlnMetAlaGluSerTyrglyLysAspArgLysArgValMet 265	
Db	14720	CCACAGCTCTGCTGCTCTATGGCTGAAGCTATTTTGAAAGACACCAACGCGTGTGTT 14661	
Qy	266	ValCysSerCysTyrglyGlnTyrglyValGlnAsnHisTyrglyLeuGlyValPro 285	
Db	14660	CCTGTGCAGCTTATCTTTACGGGCAATATGGGTCAAAGATACCTATGTTGGTGTTCGG 14601	
Qy	286	CysValIleGlyGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGlu 305	
Db	14600	GTTGTGATTGGTGGGGGGGTTGAGCGCGTGAATCGATTCGATTAAGAGGAG 14541	

Qy	306	ArgGlnGlnLeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaAla 325	
Db	14540	AAAGCTGCTTTTGAAAAATCAGTGAGTGCAGTGCACAAAACCTTGTGAAGCTGTATTGCT 14481	
Qy	326	Leu 326	
Db	14480	GTT 14478	
RESULT 14			
AE009456			
LOCUS			
DEFINITION	AE009456 11111 bp DNA linear BCT 20-MAR-2003		
	Brucella melitensis 16M chromosome I, section 13 of 195 of the		
	complete sequence.		
ACCESSION	AE009456 AE008917		
VERSION	AE009456.1 GI:117982006		
KEYWORDS			
SOURCE	Brucella melitensis 16M		
ORGANISM	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;		
	Brucellaceae; Brucella.		
REFERENCE	1 (bases 1 to 11111)		
AUTHORS	DelVecchio,V.G., Kapatral,V., Redkar,R.J., Patra,G., Mujer,C.,		
	Los,T., Ivanova,N., Anderson,I., Bhattacharyya,A., Lykidis,A.,		
	Reznik,G., Jablonski,L., Larsen,N., D'Souza,M., Bernal,A.,		
	Mazur,M., Goltzman,E., Selkov,E., Elzer,P.H., Hagius,S.,		
	O'Callaghan,D., Letesson,J.-J., Hasekorn,R., Kyripides,N. and		
	Overbeek,R.		
TITLE	The genome sequence of the facultative intracellular pathogen		
JOURNAL	Brucella melitensis		
PUBMED	Proc. Natl. Acad. Sci. U.S.A. 99 (1), 443-448 (2002)		
REFERENCE	11756688		
AUTHORS	2 (bases 1 to 11111)		
TITLE	DelVecchio,V.G., Redkar,R.J., Patra,G. and Mujer,C.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (13-NOV-2001) Institute of Molecular Biology and		
AUTHORS	Medicine, University of Scranton, Scranton, PA 18510, USA		
TITLE	3 (bases 1 to 11111)		
JOURNAL	Direct Submission		
REFERENCE	Submitted (13-NOV-2001) Department of Veterinary Science, LSU Ag		
AUTHORS	Center, 111 Dalrymple Building, Baton Rouge, LA 70803, USA		
TITLE	4 (bases 1 to 11111)		
JOURNAL	Kapatral,V., Los,T., Ivanova,N., Anderson,I., Bhattacharyya,A.,		
REFERENCE	Lykidis,A., Reznik,G., Jablonski,L., Larsen,N., D'Souza,M.,		
AUTHORS	Bernal,A., Mazur,M., Goltzman,E., Selkov,E., Hasekorn,R.,		
TITLE	Kyripides,N. and Overbeek,R.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (13-NOV-2001) Integrated Genomics, Inc., 2201 W. Campbell		
AUTHORS	Park Drive, IL 60612, USA		
TITLE	5 (bases 1 to 11111)		
JOURNAL	Letesson,J.-J.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (13-NOV-2001) Unite de Recherche en Biologie Moleculaire,		
TITLE	Laboratoire d'Immunologie et de Microbiologie, Universite of Namur,		
JOURNAL	61 rue de Bruxelles, Namur 5000, Belgium		
REFERENCE	6 (bases 1 to 11111)		
AUTHORS	O'Callaghan,D.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-NOV-2001) Faculte de Medecine, INSERM U431, Avenue		
FEATURES	Kennedy, Nimes 30900, France		
source	1. .11111		
	/organism="Brucella melitensis 16M"		
	/mol_type="genomic DNA"		
	/strain="16M"		
	/db_xref="taxon:224914"		
	/chromosomes="I"		
	complement(74..847)		
	/genes="BMEI0129"		
	complement(74..847)		
	/genes="BMEI0129"		
	/EC_number="3.1.2.6"		
	/codon_start=1		
gene			
CDS			

```

/transl_table=11
/product="hydroxacylglutathione hydrolase"
/protein_id="AAL51311.1"
/db_xref="GI:17982007"
/translation="MEORLEIEQICRSNDYGVLIHDPESALTATIDAPAYAEAL
ERRGKLDIFIETHHLDHVGNEPLKEKGVSIIGPEAEKAKIPGIDRTVKGDEFT
EGLKVKVISPISGHTAGSISYLLPDAKVVF7GDIPLFALCGRLFEPTPATWFLSEKL
VALFGDIALYCGHYTQNNARFALTIDPDNSALKERAKETIARLARHERMTLPSTIALE
MATNPFURHDTIRARLQDAPDEAVFABIRKRMF"
/gene
CDS
/gene="BMEI0130"
complement (880..1482)
/gene="BMEI0130"
complement (880..1482)
/transl_table=11
/product="INTRACELLULAR SEPTATION PROTEIN"
/protein_id="AAL51312.1"
/db_xref="GI:17982008"
/translation="MPPLLKALELPLLVFFFNARGEMLIERFPLGSGAPFLA
TALFMAATVIALAISWNRITLAIMPLVSGIVLVFGALTLWLHNDTFPIKMKPTIVAT
LFGGILGGGLFFGKSLGIVFDSAPFLDAEGWRKLTIRWGLFFIFLAIVNEIVNRFS
TDTWSPKVMGIPITVITLQMLPIQKHSLTDEENTAS"
complement (1542..2486)
/gene="BMEI0131"
complement (1542..2486)
/gene="BMEI0131"
complement (1542..2486)
/transl_table=11
/product="SIGNAL RECOGNITION PARTICLE RECEPTOR FTSY"
/protein_id="AAL51313.1"
/db_xref="GI:17982009"
/translation="MPAEVAPPEPKLSWFERLRRLRGLARSSSLSDSIIGGIFFTKRLDDD
TLQLELDVIOADLGTAKRTVTDALASGHYKDVGPGEVRAIMSAIEIKVLAPVAKP
LELDLHKPHVILVGVNGTKTTTIGKLAALKTAGGLKVMLAGDTPFAAAIBOLHI
WGETSGSPVSKLGADAGLADANWEKAKESDVLIIIDTAGRLQNKAEMLDELAKI
VRVUGKHDPAPHVITVQLDITATQCNALNQVEIFKNVAGVNGLVMTKLDGTARGGILV
AUSARHKLPIYIVGVSGIDPLEPFAAKOFARATAGVA"
complement (2961..4304)
/gene="BMEI0132"
complement (2961..4304)
/gene="BMEI0132"
complement (2961..4304)
/EC_number="1.8.-.-"
/transl_table=11
/product="Fe-S OXIDOREDUCTASE"
/protein_id="AAL51314.1"
/db_xref="GI:17982010"
/translation="MEPRHAGFAGRRORGGCLMSVEIVTFGRLNTYSEVMKREA
DAAGLGLTKGAILFNCAVTAFAEVARQARQAIKARENPDARIIVTGCAAQTEADNF
AAMEGVLDLVLEEKLSNSYRMLPDFGQFQKRVNDIMEVRETAHSHVMDAIEGRA
RAFVQVQNCCHRTCTFCII PYGRGNSRSPVMAVVDQVKRLVNGYAEVVLTVGDWTS
YGPLPSNLRLGLVKTVLAQVPLQRLRLSSIDSI EADEDLMEAIASEKRLPHLHL
SLOAGDDMLKMRKRLHRLDSDIRFCQTARALRPDI VFGADIINGFTTETEEFPQNSL
KIVBECGDLHLHVPFSPREGTTPAARMPQVRREIVKERAARLRABGRAYEKKHLNALH
GTROLLIVEKEGVARTGFTLAAVDQGNAGEIIRIIVTHGDGKLLTRQAPQAA"
complement (4244..5155)
/gene="BMEI0133"
complement (4244..5155)
/gene="BMEI0133"
complement (4244..5155)
/EC_number="5.1.1.7"
/transl_table=11
/product="DIAMINOPIMELATE EPIMERASE"
/protein_id="AAL51315.1"
/db_xref="GI:17982011"
/translation="MATKAAPARMNGLNQIIVADMGRADSIITSAAIIRLASDETA
FDQIMATHDPTGTDTYIAIINCDDGTQAQACNGTRCVVQALAAETGRHAFPTETA
GILFATHDHGLISVDMDTPREDWDQIPLAQAVADTRKIELOVDPADPADVLHSPSIAS
MGNHFAVVDKQVMSYELDKFGLLENHPIFPERANISIAHVTSSDITLRTWGERA
GLTRACSGAACAAVARSRTGRKVTNVVPGGPLLIEWRDDDDHHVMTGPAWEFSG
TFDPATGEWGRDRTQGLQSGNADRGAA"
5186..5389
/gene="BMEI0134"

```

```

CDS
5186..5389
/gene="BMEI0134"
/codon_start=1
/transl_table=11
/product="Hypothetical Protein"
/protein_id="AAL51316.1"
/db_xref="GI:17982012"
/translation="MVIKPDNEGLYGGKARKLPVPARDCNLCQSHAPRTNANIGETV
AFICQFQNYQVQVSSILEGLRQK"
5518..6051
/gene="BMEI0135"
5518..6051
/gene="BMEI0135"
/codon_start=1
/transl_table=11
/product="OUTER MEMBRANE LIPOPROTEIN"
/protein_id="AAL51317.1"
/db_xref="GI:17982013"
/translation="MGISKASLLSLAAGIVLAGCQSSRLGNLDNVSPPPPPAPVNAV
PACTVQGNLSDPTQFPNAPSTMSAQSGTQVASLPASAPDLTPGAVAGVWNASLGG
QSKKIATPQTKYGGYRAGPLRCPGELANLASWAVNGKQLVLYDANGGTVASLYSSQG
GRFDGOTTGGQAVTLGR"
6123..7334
/gene="BMEI0136"
6123..7334
/gene="BMEI0136"
/codon_start=1
/transl_table=11
/product="PUTATIVE ATPASE N2B"
/protein_id="AAL51318.1"
/db_xref="GI:17982014"
/translation="MISADAHQSKVSLPLEPLMSFDNLKAPFSVRERYEAMVAAGDVT
DPAQLELTSDYRLIEIEICSKRLSRSSALGWLFGKRRETAATIKGLYVHGEVGRGT
MLMDQFQLLPVERKRAHENDMADVHERIYAHROAKHGETKQDDPIPPVAEALSO
QAVLCPDEFETVIDIADAMILSRFLSALFSRGVVVLVATSNVAPDNLVRDGLNRQLFLP
FIDILKSHVDVINLSDRTDYRLKLDQPVLSPLGPETERRMDAAWAHKNKGAEEKP
DVTHIKGRDIEVPRAAGAAARFDDLCARPLGASDYIAIANRYPTLFIQNVPLVDYS
RRNEAKRFILLIDVLDYDHARLFVSAEAQPEKLYIANSGTAEAFEDRTASRLFEMQSA
EYLDGIPEGVA"
7755..8735
/gene="BMEI0137"
7755..8735
/gene="BMEI0137"
/EC_number="1.1.1.37"
/codon_start=1
/transl_table=11
/product="MALATE DEHYDROGENASE"
/protein_id="AAL51319.1"
/db_xref="GI:17982015"
/translation="MRKETIMARKNIALIGSGMIGGTLAHLAGLKGLDVVLFDIAEG
TPQKGLDIAESSPVDGFDAKFTGANDYAAIEGADVVIITAGVPRKPGMSRDDLLGIN
LKTMEQVGAGIKKYAREAFVICITNPLDMVWALQKESGLPAHKVIVGMAGVLDASFR
YFUSEFNVSVEDVTAFVLGGHGSWVPLARYSTVAGIPLSLDLVKGWNTSQDKLDKII
QRTRDGAAEIVGLLKTGSAPYAPAAIAIQMAESYLLDKKRVLPVAAQLSGQYGVDMY
VGVTYVIGANGVERIIEIDLKDEKAQFDKSVASVAGLCEACIGIAPSLK"
8842..10038
/gene="BMEI0138"
8842..10038
/gene="BMEI0138"
/EC_number="6.2.1.5"
/codon_start=1
/transl_table=11
/product="SUCCINYL-COA SYNTHETASE BETA CHAIN"
/protein_id="AAL51320.1"

```

Alignment Scores:

Pred. No.:	1,21e-63	Length:	11111
Score:	876.50	Matches:	174
Percent Similarity:	70.7%	Conservative:	53
Best Local Similarity:	54.2%	Mismatches:	89
Query Match:	51.9%	Indels:	5
DB:	1	Gaps:	1

US-09-390-846-2 (1-330) x AE009456 (1-11111)		
Qy	9	ArgProlysisIleAlaMetValGlySerGlyMetIleGlyGlyThrMetAlaPheLeuCys 28
Db	7779	CGCAACAGATTGGCTTCGGCTCGCGCATGATCGCGGTACGCTCGCTCACTCGCC 7838
Qy	29	SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMetGly 48
Db	7839	GGTCTGAAGGAACCTGGTACGCTGCTTTTCACATTCGGGAAGGCACCCCGCAGGC 7898
Qy	49	LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyrGly 68
Db	7899	AAGGGCTGGATATCGCGGAATCTCTCCGTCGATGTTTCGATGCGAAGTTCACCTGGC 7958
Qy	69	SerAsnSerTyrGluCysLeuGlyValAlaAspValValIleIleThrAlaGlyIleThr 88
Db	7959	CGCAACGATTACGGCGCATTAAGAGCGGCACACGCTCGTCATCGCACGGGGTCCG 8018
Qy	89	LysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProValAsnIle 108
Db	8019	CGCAAGCGGCG-----ATGAGCGCGCAGCATCTCTCGGGCATCAACTG 8063
Qy	109	LysIleMetArgGluValGlyAlaAlaIleLysSerTyrCysProAsnAlaPheValIle 128
Db	8064	AAGGTGATGGAACAGGTGGCGCGGCATCAAGAAATATGCCCGGAAGCTTTCGTCTATC 8123
Qy	129	AsnIleThrAsnProLeuAspValMetValAlaAlaLeuGlnGluSerSerGlyLeuPro 148
Db	8124	TGCATCAACCAACCGCTCGACGCATGCTTGGCGCTGTCAGAAAGTTCCTCGGCTTCCG 8183
Qy	149	HisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMetIle 168
Db	8184	CCCCACAGGTTCGGCATGCTGGCGTCTCCACAGCGCCGCTTCGTTATTTCTC 8243
Qy	169	AlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHisGly 188
Db	8244	TCGGGAAGAAATCAACGTTTCGGTCGAGACGTCACGGCATTCGTCGGTGGCCACGGC 8303
Qy	189	AspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGluPhe 208
Db	8304	GATTTCGATGGTTCCGCTGGCGCGCTACTCGACCGTTCGCGCATTTCCGCTGCTCGATCTC 8363
Qy	209	ValLysLysGlyTyrIleLysGlnGluValAspIleValGlnLysThrLysVal 228
Db	8364	GTCAGATGGCTGGACCGCAGCAGCAAGCTCCACAGATCATCCAGCGCACCGGTGAC 8423
Qy	229	AlaGlyGlyGluIleValArgLeuGlyGlnGlySerAlaTyrTyrAlaProGlyAla 248
Db	8424	GGCGCGCGGAAATCGTGGGCTTCTCAAGACCGGCTCGGCTTCTAGCTTCGCGCGCA 8483
Qy	249	SerAlaIleGlnMetAlaGluSerTyrLeuLysAspArgLysArgValMetValCysSer 268
Db	8484	TCGGCCATCCAGATGGCTGGAATCTTACCTCAAGGACAAAGACGCGCTCTCGCGGTGCGCA 8543
Qy	269	CysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysValIle 288
Db	8544	CCCCAGCTTCGGGCCAGATATGGGTGAAGACATGATGTGGCGGTTCGACCGTATC 8603
Qy	289	GlyGlyArgGlyValGluLysIleIleGluLeuGluLeuThrAlaGlnGluArgGlnGlu 308
Db	8604	GGCGCAATGGCTGGAGCGCATCATCGAGATCGATCTCGACAAGGACGAGAGAGCGCAG 8663
Qy	309	LeuGlnGlySerIleAspGluValLysGluMetGlnLysAlaIleAlaLeuAspAla 328
Db	8664	TTCGACAAGTCGGTGGCTCGGCTCGCGGTCTGTGCGAAGCCTGCATCGGTATTGCTCCG 8723
Qy	329	Ser 329
Db	8724	TCG 8726
RESULT 15		
AF322647		
LOCUS		963 bp DNA linear BCT 11-FEB-2005

US-09-390-846-2 (1-330) x AF322647 (1-963)		
Qy	9	ArgProlysisIleAlaMetValGlySerGlyMetIleGlyGlyThrMetAlaPheLeuCys 28
Db	7	CGCAACAGATTCGACATTATCGGTTTCGGGATGATTGGTGGCAGCTGGCGCATCTCGCC 66
Qy	29	SerLeuArgGluLeuGlyAspValValLeuPheAspValValProAsnMetProMetGly 48
Db	67	GGCTGAAGGAACCTGGCGACATCGCTCTCTCGACATTCGCCGACGCGCATTCGCCAGGC 126
Qy	49	LysAlaMetAspIleSerHisAsnSerSerValValAspThrGlyIleThrValTyrGly 68
Db	127	AAGGCTTCGACATCGCGAGTCTCTCCCGTCGAGGGCTTCGATGTCATCGCTGACCGGT 186
Qy	69	SerAsnSerTyrGluCysLeuLysGlyAlaAspValValIleIleThrAlaGlyIleThr 88
Db	187	GCAGGCGACTATTTCGCGATTCGAGGGCGCGAGCTCTGCATCTGCTACCGCGCGGTCGC 246
Qy	89	LysIleProGlyLysSerAspLysGluTrpSerArgMetAspLeuLeuProValAsnIle 108
ORIGIN		
Alignment Scores:		
Pred. No.:		1e-64
Score:		873.50
Percent Similarity:		73.2%
Best Local Similarity:		53.2%
Query Match:		51.7%
DB:		1
US-09-390-846-2 (1-330) x AF322647 (1-963)		
DEFINITION		Sinorhizobium meliloti malate dehydrogenase (mdh) gene, complete cds.
ACCESSION		AF322647
VERSION		AF322647.1 GI:11935163
KEYWORDS		Sinorhizobium meliloti (Rhizobium meliloti)
SOURCE		Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Sinorhizobium.
ORGANISM		1 (bases 1 to 963)
REFERENCE		Dymov,S.I., Meek,D.J., Steven,B. and Driscoll,B.T.
AUTHORS		Insertion of transposon TnStaci in the Sinorhizobium meliloti
TITLE		malate dehydrogenase (mdh) gene results in conditional polar
JOURNAL		effects on downstream TCA cycle genes
PUBMED		Mol. Plant Microbe Interact. 17 (12), 1318-1327 (2004)
REFERENCE		15597737
AUTHORS		2 (bases 1 to 963)
TITLE		Dymov,S.I., Meek,D.J. and Driscoll,B.T.
JOURNAL		Direct Submission
FEATURES		Submitted (20-NOV-2000) Natural Resource Sciences, McGill University, 21,111 Lakeshore Rd, Ste-Anne-de-Bellevue, Quebec H9X 3V9, Canada
source		Location/Qualifiers
gene		1..963
CDS		/organism="Sinorhizobium meliloti"
		/mol_type="genomic DNA"
		/db_xref="taxon:382"
		1..963
		/genes="mdh"
		1..963
		/genes="mdh"
		/EC_number="1.1.1.37"
		/codon_start=1
		/transl_table=11
		/product="malate dehydrogenase"
		/protein_id="AAG41996.1"
		/db_xref="GI:11935164"
		/translation="MARNKIALIGSGMIGTTLAHLAGLKLGDIVLFDIADGIPQKGKLDIAOSSPVEGPDASLTGASDYSATIEGADVCIATGVPKPKGMRDLDLGINLKVMEOVGAGIKKYPANAFVICITNPLDMWWALQKFSGLPKNVGVNAGVLDSSRFLFLAEENFVSKDITAFVLGGHGTWPLARYSTVAGIPLDLIOGWTTTKEKLDIOIORTDGGABIVGLTKTSAYAPAAASIAEMAEAYLKDKRVLPCAAHLSSGGYGVKDMYGVVPTVIGAGGTERIIEIDLNKKEAFDKSAVAAGLCEACINAPSLK"

```

Db      247 CGCAAGCCCGGC-----ATGAGCCGCGAGCATCTGCTCGGCATCAACCTC 291
Qy      109 LysIleMetArgGluValGlyAlaIleLysSerTyrCysProAsnAlaPheValIle 128
Db      292 AAGGTCATGGAACAGGTGCGCGCGGCATCAAGAAATAGCTCCGAAGCCTTCGTGATC 351
Qy      129 AsnIleThrAsnProLeuAspValMetValAlaIleLeuGlnGluSerSerGlyLeuPro 148
Db      352 TGCATCACCAACCCCTGACCCATGGTCTGGGCGCTGCGAAGATTCTTCGGCCCTGCGG 411
Qy      149 HisHisArgIleCysGlyMetAlaGlyMetLeuAspSerSerArgPheArgMetIle 168
Db      412 AAGAACAAGGTGTCGCGCATGCGCGCGTCTCGACTCTTCGGCTTCGGCTCTTCCTC 471
Qy      169 AlaAspLysLeuGluValSerProArgAspValGlnGlyMetValIleGlyValHisGly 188
Db      472 GCCGAAGAGTTCAACGTCTCTGTCGAAGACATCACCGCTTCGTGCTCGGGCGCATGGC 531
Qy      189 AspHisMetValProLeuSerArgTyrAlaThrValAsnGlyIleProLeuSerGluPhe 208
Db      532 GACACCATGGTCCGCTCGCGGTATTTCGACGGTTGCGGCGATCCCGTGCCTGCGACCTG 591
Qy      209 ValLysLysGlyTyrIleLysGlnGluValAspAspIleValGlnLysThrLysVal 228
Db      592 ATCCAGATGGTTGACCAACCAGAAAGCTCGACCATTCATTCAGCGCACCGGTGAC 651
Qy      229 AlaGlyGlyGluIleValArgLeuLeuGlyGlnGlySerAlaTyrTyrAlaProGlyAla 248
Db      652 GCGCGCGCGAGATCGTCGGCTTGCTCAAGACCGCTCGGCTATTATGCGCGCGCGCC 711
Qy      249 SerAlaIleGlnMetAlaGluSerTyrIleLysAspArgLysArgValMetValCysSer 268
Db      712 TCGGCGATCGAAATGCCGAAGCCTATCTCAAGGACAAGACCGCGCTGCTGCTGCGCC 771
Qy      269 CysTyrLeuGlnGlyGlnTyrGlyValGlnAsnHisTyrLeuGlyValProCysValIle 288
Db      772 GCTCATCTCTCGGCGCAGTATGGCGTCAAGGACATGTATGTGCGCGTCCCGCATCGTATC 831
Qy      289 GlyGlyArgGlyValGlyLysIleIleGluLeuGluLeuThrAlaGlnGluArgGlnGlu 308
Db      832 GGTGCGCGCGCATCGAGCGCATCATCGAGATCGACCTCAACAAGGGCGGAGAGGAGCC 891
Qy      309 LeuGlnGlySerIleAspGluValLysGluMetGlnLysAla 322
Db      892 TTCGACAAAGTCGGTTGCGCGCGCTCGCGCGCTTCGCGAAGCC 933

```

Search completed: March 3, 2006, 01:15:51
Job time : 4623 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 19:41:36 ; Search time 67 Seconds
(without alignments)
2057.965 Million cell updates/sec

Title: US-09-390-846-2
Perfect score: 1688
Sequence: 1 MAVFKNTRPKIAMVSGMI.....GSIDEVEMQKATAALDASK 330

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA_Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	880.5	52.2	320	4	US-10-369-493-17838
2	873.5	51.7	319	4	US-10-369-493-17838
3	873.5	51.7	320	4	US-10-369-493-17838
4	873.5	51.7	320	4	US-10-369-493-17838
5	862.5	51.1	311	4	US-10-369-493-17838
6	862.5	51.1	322	4	US-10-369-493-17838
7	850.5	50.4	320	4	US-10-369-493-17838
8	848.5	50.3	320	4	US-10-369-493-17838
9	847.5	50.2	320	4	US-10-369-493-17838
10	740	43.8	312	4	US-10-369-493-17838
11	709	42.0	314	4	US-10-369-493-17838
12	697	41.3	312	4	US-10-369-493-17838
13	689	40.8	285	4	US-10-369-493-17838
14	672	39.8	309	4	US-10-369-493-17838
15	669.5	39.7	319	4	US-10-369-493-17838
16	652.5	38.7	304	4	US-10-369-493-17838
17	649.5	38.5	325	4	US-10-369-493-17838
18	641	38.0	325	4	US-10-724-972A-7257
19	635	37.6	304	4	US-10-369-493-17838
20	635	37.6	317	4	US-10-369-493-17838
21	634	37.6	324	4	US-10-369-493-17838
22	628	37.2	299	4	US-10-369-493-17838
23	626.5	37.1	305	4	US-10-369-493-17838
24	625.5	37.1	309	4	US-10-369-493-17838
25	605.5	35.9	335	4	US-10-369-493-17838
26	569	33.7	334	4	US-10-369-493-17838
27	560.5	33.2	313	4	US-10-369-493-17838

28	529	31.3	197	4	US-10-369-493-17838
29	478	28.3	314	4	US-10-369-493-17838
30	462	27.4	312	4	US-10-369-493-17838
31	460	27.3	321	4	US-10-369-493-17838
32	458	27.1	312	4	US-10-369-493-17838
33	455	27.0	318	4	US-10-369-493-17838
34	439.5	26.0	319	4	US-10-369-493-17838
35	439	26.0	312	4	US-10-369-493-17838
36	436	25.8	319	3	US-09-971-361-9
37	436	25.8	319	3	US-09-971-361-9
38	436	25.8	319	6	US-11-108-870-3
39	436	25.8	319	6	US-11-108-870-3
40	425	25.2	318	4	US-10-369-493-17838
41	422	25.0	308	4	US-10-369-493-17838
42	418.5	24.8	323	5	US-10-873-467-76
43	411.5	24.4	310	4	US-10-369-493-17838
44	408	24.2	332	4	US-10-170-385-223
45	408	24.2	332	5	US-10-873-595-3

ALIGNMENTS

RESULT 1
US-10-369-493-17838
; Sequence 17838, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17838
; LENGTH: 320
; TYPE: PRT
; ORGANISM: SPHINGOMONAS
US-10-369-493-17838

Query Match	52.2%	Score	880.5	DB	4	Length	320
Best Local Similarity	53.9%	Pred. No.	2e-79				
Matches	173	Conservative	54	Mismatches	89	Indels	5
Gaps	1						
QY	9	RPKIAMVSGMIGTMAFLCSLRELGDVVLFDVVPNMPGKAMDISHNSSVVDGTITVG	68				
Db	3	RKIALIGSGMIGTMAFLCSLRELGDVVLFDVVPNMPGKAMDISHNSSVVDGTITVG	62				
QY	69	SNSYECLGADVVIIITAGITKIPKSDKEMSRMDLLFVNIKIMREYGAATKSCYCPNAFVI	128				
Db	63	ANSYEDIAGDVCIIVTAGIPRKEG-----MSRDDLLKTNLGVKMKVAGEGIAAHAPDAFVI	117				
QY	129	NITNPLDVMVAALQSSGLPHHRCIMAGMDSRRFRMIADKLEVSPRDVQGVIGVHG	188				
Db	118	CITNPLDVMVAALREFSGLPHQKVMGAGVLDARSFRHFLAEFNFVSVEDVTFVGLGGHG	177				
QY	189	DHMPVLSRYATVNGIPLSEFVKKGWKQREVDVIVOKTVKAGGEIVRLLGQSGAYVAPGA	248				
Db	178	DTNPVIEYSTVAGIVPDLIRKMGWSTQBRIDAIVARTSGGGEIVALLTKTSAYYAPAT	237				
QY	249	SAIQMAESYLKDKRKRVWVGVSCVYQGVQVNHVYLPVPCVIGGVRGVEKIIIELELTACERQE	308				
Db	238	SALMAESYLKDKRLLPCCAHLTGQGVDDLLVGVPIVIGKDGVERIVEIELNATAKQN	297				
QY	309	LOGSIDEVEMQKATAALDAS	329				

Sequence	14326	A
Sequence	18330	A
Sequence	16679	A
Sequence	23073	A
Sequence	16703	A
Sequence	46	Appl
Sequence	3069	Ap
Sequence	16476	A
Sequence	3	Appl
Sequence	9	Appl
Sequence	3	Appl
Sequence	3022	Ap
Sequence	3200	Ap
Sequence	76	Appl
Sequence	17474	A
Sequence	223	App
Sequence	3	Appl

```

; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14489
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-14489

Query Match          51.7%; Score 873.5; DB 4; Length 320;
Best Local Similarity 53.2%; Pred. No. 1e-78;
Matches 167; Conservative 60; Mismatches 82; Indels 5; Gaps 1;

QY 9 RPKIAMVGSIGMGTMAFLCSLRELGDVLFVDPVNNPMGKAMDISHNSSVVDGTITVYG 68
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 3 RKKIALIGSGMIGTTLAHLASLKLKELGDIVLFDIADGIPQGGKGLDIAQSGPVEGFNAKLSG 62
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 69 SNSYECLKGADVVIITAGITKIPGSKDKWSRMDLLPVNIKIMREVGAAIKSYCPNAPVI 128
    :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 63 ASDYAAIEGADVCIIVTAGVARKPG-----MSRDDLLGINLKWMEQVGAGIKKYAPNAPVI 117
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 129 NITNPLDMVMAALQESSGLPHHRICGMAGMLDSSRRFRMIADKLEVSPRDVGQWVGIVHG 188
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 118 CITNPLDMVMAALQKFSGLPKNVGMAGVLDLSARFRLPLAEEFNVSQDVTAFVLGGHG 177
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 189 DHMVPLSRVATNGIPLSEFVKKGWIKQEEVDIVQKTKVAGGEIVRLQGSAYYAPGA 248
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 178 DTWVPLARYSTVGVPVLTDLVKMGWLTAELEQIIQTRDGGAEIVGLLKTGSAYYAPAA 237
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 249 SAIOAESYLKDKRVMVCSYLOGQYGVQNHVLPVCPVIGGRGVEKIIIELELTAQERQE 308
    :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 238 SAIEMAESYLKDKRVLPAHAHLSGQYGVDDMVGVPTIIGAGGIERVIEIENKEEEAA 297
    :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 309 LQGSIDEVKEMQKA 322
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 298 FOKSVGAVAGLCSEA 311
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 4
US-10-369-493-14874
; Sequence 14874, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14874
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-14874

Query Match          51.7%; Score 873.5; DB 4; Length 320;
Best Local Similarity 53.2%; Pred. No. 1e-78;
Matches 167; Conservative 60; Mismatches 82; Indels 5; Gaps 1;

QY 9 RPKIAMVGSIGMGTMAFLCSLRELGDVLFVDPVNNPMGKAMDISHNSSVVDGTITVYG 68
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 3 RKKIALIGSGMIGTTLAHLASLKLKELGDIVLFDIADGIPQGGKGLDIAQSGPVEGFNAKLSG 62
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 69 SNSYECLKGADVVIITAGITKIPGSKDKWSRMDLLPVNIKIMREVGAAIKSYCPNAPVI 128
    :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 63 ASDYAAIEGADVCIIVTAGVARKPG-----MSRDDLLGINLKWMEQVGAGIKKYAPNAPVI 117
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 129 NITNPLDMVMAALQESSGLPHHRICGMAGMLDSSRRFRMIADKLEVSPRDVGQWVGIVHG 188
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 118 CITNPLDMVMAALQKFSGLPKNVGMAGVLDLSARFRLPLAEEFNVSQDVTAFVLGGHG 177
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 189 DHMVPLSRVATNGIPLSEFVKKGWIKQEEVDIVQKTKVAGGEIVRLQGSAYYAPGA 248
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 178 DTWVPLARYSTVGVPVLTDLVKMGWLTAELEQIIQTRDGGAEIVGLLKTGSAYYAPAA 237
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 249 SAIOAESYLKDKRVMVCSYLOGQYGVQNHVLPVCPVIGGRGVEKIIIELELTAQERQE 308
    :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 238 SAIEMAESYLKDKRVLPAHAHLSGQYGVDDMVGVPTIIGAGGIERVIEIENKEEEAA 297
    :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 309 LQGSIDEVKEMQKA 322
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 298 FOKSVGAVAGLCSEA 311
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 3
US-10-369-493-14489
; Sequence 14489, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21

```

Db 118 CTTNPLDAMWALQKFGSLPKKVKVGMAGVLDLSARFLFLAEENFVNSVDVTFVLGGHG 177
Qy 189 DHMVLPSRYATVNGIPSEFVKKWKIKOEVDVDDIVQKTKVAGGIVRLLLQGSAYYAPGA 248
Db 178 DTWVPLARYSTVGGVPLTDLVKGWLTAELEQIQRTRDGGAEIVGLLTKGSAYYAPAA 237
Qy 249 SATQMAESYLKDKRVMVCSYLOGQYGVQNHVGLVPCVIGRGVKEKIELELTAQERQE 308
Db 238 SALEMAESYLKDKRVLPAHAHLGSGQYGVDDMYGVVPTIIGAGGIERVIELNKEBEAA 297
Qy 309 LQGSIDEVKEMQKA 322
Db 298 FQKSVGAVGLCEA 311

RESULT 5

US-10-369-493-20748
; Sequence 20748, Application US/10369493
; Publication No. US20030233675A1

GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 20748
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Rhodospseudomonas palustris
US-10-369-493-20748

Query Match 51.1%; Score 862.5; DB 4; Length 311;
Best Local Similarity 53.8%; Pred. No. 1.3e-77;
Matches 169; Conservative 59; Mismatches 81; Indels 5; Gaps 1;
Qy 9 RPKIAMVGSOMIGTMAFLCSLRELGDVVLFDVVPNNPMKAMDISHNSSVVDITGITYG 68
Db 3 RDKIALIGSGIGTTLAHLVGLKELGDVVLFDIAEGVPGKALDIAESSPVDGFDKLTG 62
Qy 69 SNSYECLKGADVVIITAGITKIPKSKDKEWSDMLLPVNIKIMREVGAAIKSCYCPNAFVI 128
Db 63 ANSYEATEGARVVIVTAGVPRKPG-----MSRDDLLSINLKVMEQVAGIKKYAPDAFVI 117
Qy 129 NITNPLDVMVAALQESSGLPHRRICGMAGMLDSRFRMIADKLEYSPRDVQGMVIGVHG 188
Db 118 CTTNPLDAMWALQKAGSLPAKKVGMAGVLDLSARFLFLAEENFVNSVDVTFVLGGHG 177
Qy 189 DHMVLPSRYATVNGIPSEFVKKWKIKOEVDVDDIVQKTKVAGGIVRLLLQGSAYYAPGA 248
Db 178 DTWVPLVKYSTVAGIPLDVLKMGWTSQARLDELIVDTNRNGAEIVNLLTKGSAYYAPAS 237
Qy 249 SATQMAESYLKDKRVMVCSYLOGQYGVQNHVGLVPCVIGRGVKEKIELELTAQERQE 308
Db 238 SALEMAESYLKDKRVPVAAHLNGEYGVKDMYGVVPVWIGDKGVERIVEIELAGKDEA 297
Qy 309 LQGSIDEVKEMQKA 322
Db 298 FDRSVAAVQGLVEA 311

RESULT 6

US-10-369-493-12060
; Sequence 12060, Application US/10369493
; Publication No. US20030233675A1

GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 12060
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Mesorhizobium loti
US-10-369-493-12060

Query Match 51.1%; Score 862; DB 4; Length 322;
Best Local Similarity 53.5%; Pred. No. 1.5e-77;
Matches 174; Conservative 56; Mismatches 87; Indels 8; Gaps 2;
Qy 9 RPKIAMVGSOMIGTMAFLCSLRELGDVVLFDVVPNNPMKAMDISHNSSVVDITGITYG 68
Db 3 RNKIALIGSGIGTTLAHLVGLKELGDVVLFDIAEGVPGKGLDIAQSSPVDGFDKLTG 62
Qy 69 SNSYECLKGADVVIITAGITKIPKSKDKEWSDMLLPVNIKIMREVGAAIKSCYCPNAFVI 128
Db 63 VNDYAGIEGADVIVTAGVPRKEG-----MSRDDLLSINLKVMEQVAGLUKKTAPKAFVI 117
Qy 129 NITNPLDVMVAALQESSGLPHRRICGMAGMLDSRFRMIADKLEYSPRDVQGMVIGVHG 188
Db 118 CTTNPLDAMWALQKSGSLPKTHVGMAGVLDLSARFLFLAEENFVNSVDVTFVLGGHG 177
Qy 189 DHMVLPSRYATVNGIPSEFVKKWKIKOEVDVDDIVQKTKVAGGIVRLLLQGSAYYAPGA 248
Db 178 DSMVPMIRYSTVSGIPLDVLKMGWTSKEKLDQIVQTRDGGAEIVGLLTKGSAYYAPAA 237
Qy 249 SATQMAESYLKDKRVMVCSYLOGQYGVQNHVGLVPCVIGRGVKEKIELELTAQERQE 308
Db 238 SALEMAESYLKDKRVLPCAHLGSGYGVKGTIVGVVPVWIGAGGVERIIEIDLNKSEQKM 297
Qy 309 LQGSIDEVKEMQKA-----IAALDASK 330
Db 298 FESSVATVQGLTEACVKIAPQLASK 322

RESULT 7

US-10-369-493-17079
; Sequence 17079, Application US/10369493
; Publication No. US20030233675A1

GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17079
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Caulobacter crescentus
US-10-369-493-17079

; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 23237

; LENGTH: 312

; TYPE: PRT

; ORGANISM: Bacillus subtilis

US-10-369-493-23237

Query Match 43.8%; Score 740; DB 4; Length 312;

Best Local Similarity 45.9%; Pred. No. 2.5e-65;

Matches 147; Conservative 63; Mismatches 98; Indels 12; Gaps 3;

Qy 7 NTRPKIAMVSGMIGGTMAFLCSRELGDVVLFDV--VPNMPMGKAMDISHNSSVVDGTI 64

Db 3 NTRKKSIVGAGTGATTAFLIAQKELADVLDIPOLENPTKGKALDMLLEASPVOGFDA 62

Qy 65 TVYGSNIECLKGADVVIIITAGITKIPKSDKEWSRMDLLPNNIKIMREVGAAIKSYCPN 124

Db 63 KITGTSNYEDTAGSDIVVITAGIARKEG-----MSRDDLVSTNEKIMRSVTQETVKYSPD 117

Qy 125 AFVINITNPLDVMVAALQESSGLPHHRICGMAGMLDSSRRFRMIADKLEVSPRDVQGMVI 184

Db 118 SIIIVLTNPVDAMTYAYKESGFFKRVIGQSGVLDRTARFTFVAEELNSVKDVTGFVL 177

Qy 185 GVHGDHMPVLSRYATVNGIPLSEFVKKGWIKQEBVDIVQKTKVAGGEIVRLLGQSAYY 244

Db 178 GGGDDMVPLIRYSYAGGIPLTLIPK-----ERIDAIVERTRKGGEIVNLLGNGSAYY 232

Qy 245 AGASAIQMAESYLKDRKRVWVCSCYLQGGVQVQNHVLPVPCVIGRGVEKIIIELELTAQ 304

Db 233 APAASLTEMVEAILKQRRVLPPTIAYLEGEYGYEIVLPTVGGNGLEQIIIELELTDY 292

Qy 305 EROELQGSIDEVKEKQAKA 324

Db 293 ERAQLNKSVESVKNVMKVL 312

RESULT 11

US-10-369-493-17395

; Sequence 17395, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 17395

; LENGTH: 314

; TYPE: PRT

; ORGANISM: Bacillus halodurans

US-10-369-493-17395

Query Match 42.0%; Score 709; DB 4; Length 314;

Best Local Similarity 44.0%; Pred. No. 3.3e-62;

Matches 140; Conservative 67; Mismatches 97; Indels 14; Gaps 4;

Qy 9 RPKIAMVSGMIGGTMAFLCSRELGDVVLFDVVPNM---PMGKAMDISHNSSVVDGTI 65

Db 5 RRVKSVIGAGTGATTALMAVQAQKELGDVLDV-IPQMEGPTKGKALDMLLESTPVQGDVN 63

Qy 66 VYGSNIECLKGADVVIIITAGITKIPKSDKEWSRMDLLPNNIKIMREVGAAIKSYCPNA 125

Db 64 ITGTSYETTKSDSDVVVITAGIARKEG-----MSRDDLVSTNAGIMKAVTKKEVVKHSPNA 118

Qy 126 FVINITNPLDVMVAALQESSGLPHHRICGMAGMLDSSRRFRMIADKLEVSPRDVQGMVIG 185

Db 119 YIIIVLTNPADAMTYTYVYKESGFFKRVIGQSGVLDRTARFTFVAQELNLSVEDITGPFVL 178

Qy 186 VGHDMVPLSRATVNGIPLSEFVKKGWIKQEBVDIVQKTKVAGGEIVRLLGQSAYYA 245

Db 179 GHGDDMVPLIRYSYAGGIPLTLIPK-----LPOERIDAIVERTRKGGEIVGLLNGSAYYA 233

Qy 246 PGASAIQMAESYLKDRKRVWVCSCYLQGGVQVQNHVLPVPCVIGRGVEKIIIELELTAQ 305

Db 234 PAASLAEMVEAILKQRRVLPPTIAYLEGEYGYEIVVGVPTILGGDGIKIELELTDDE 293

Qy 306 ROELQGSIDEVKEKQAKI 323

Db 294 KATFAKSIESVRNVSAL 311

RESULT 12

US-10-369-493-16518

; Sequence 16518, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 16518

; LENGTH: 312

; TYPE: PRT

; ORGANISM: Bacillus thuringiensis

US-10-369-493-16518

Query Match 41.3%; Score 697; DB 4; Length 312;

Best Local Similarity 43.5%; Pred. No. 5.2e-61;

Matches 138; Conservative 65; Mismatches 102; Indels 12; Gaps 3;

Qy 9 RPKIAMVSGMIGGTMAFLCSRELGDVVLFDV--VPNMPMGKAMDISHNSSVVDGTI 66

Db 5 RRVKSVIGAGTGATTTFFLLAQKELADVLDIPOLENPTKGKALDMLLEASPVOGFDA 64

Qy 67 YGSNIECLKGADVVIIITAGITKIPKSDKEWSRMDLLPNNIKIMREVGAAIKSYCPNA 126

Db 65 IGTSYADTADSDVVVITAGIARKEG-----MSRDDLVATNSKIMKSIITRDIKHSN 119

Qy 127 VINITNPLDVMVAALQESSGLPHHRICGMAGMLDSSRRFRMIADKLEVSPRDVQGMVIG 186

Db 120 IVVLTNPVDAMTYSVFKKAGFPKRVIGQSGVLDRTARFTFVAQELNLSVKDITGPFVL 179

Qy 187 HGDHMPVLSRYATVNGIPLSEFVKKGWIKQEBVDIVQKTKVAGGEIVRLLGQSAYYAP 246

Db 180 HGDDMVPLIRYSYAGGIPLTLIPK-----ERLEAIVERTRKGGEIVGLLNGSAYYAP 234

Qy 247 GASAIQMAESYLKDRKRVWVCSCYLQGGVQVQNHVLPVPCVIGRGVEKIIIELELTAQ 306

Db 235 AASLWEMTEAILKQRRVLPPTIAYLEGEYGYSDLVGLVGPVILGGNGIEKIELELTADEK 294

Qy 307 ROELQGSIDEVKEKQAKI 323

Db 295 EALDRSVESVRNVMKVL 311

RESULT 13

US-10-369-493-19127

; Sequence 19127, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 19127

; LENGTH: 285

; TYPE: PRT

; ORGANISM: Myxococcus xanthus

US-10-369-493-19127

Query Match 40.8%; Score 689; DB 4; Length 285;

Best Local Similarity 47.1%; Pred. No. 2.9e-60;

Matches 139; Conservative 58; Mismatches 86; Indels 12; Gaps 3;

QY 9 RPKIAMVSGMIGTMAFLCSLRELGDVVLFDV--VPNMPMGKAMDISHNSSVVDGTIV 66

DB 1 KKKIGLGGQIGGNLALLAVQKSLGDVLYDIPAAEGLVKGALDINQLAAVDGDCRV 60

QY 67 YGNSYECLKADVVITAGITKIPGSKDKWSRMDLLPVNIKIMREVGAIAKSYCPNAP 126

DB 61 KGTDDKDVAGSDVIIITAGMPKPG-----MSREDLLEINLKIMTIDVAGNIQKAPNAP 115

QY 127 VINITPLDVMAALQESSGLPHHRICGMAGMLDSSRFRMIADKLEVSPRDVGQVIGV 186

DB 116 VINVANPLDAMVFAHKLKIALGPKHVMAGMAGVLDTSRFRKCFVABALGCSIRDVEALVLGG 175

QY 187 HGHMVPLSRYATVNGIPLSEFVKKGWIKQEEVDIVQKTKVAGGETVRLLGQGSAYYAP 246

DB 176 HGGDMPVLRHRTVGGVPLTELAK-----DKDLAIKRTREGPELVGLYKTSYFPG 230

QY 247 GASAIOMAESYLKDRKRVWCSCYLOGQYGVQNHVLPVPCVIGRGVKEKIELEL 301

DB 231 AACAIAMAESFLQDRKRVLPAAALLEGQYINGYFFGVPGVQIGAGGVKEKHTVEL 285

RESULT 14

US-10-369-493-9045

; Sequence 9045, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 9045

; LENGTH: 309

; TYPE: PRT

; ORGANISM: Chloroflexus aurantiacus

US-10-369-493-9045

Query Match 39.8%; Score 672; DB 4; Length 309;

Best Local Similarity 43.8%; Pred. No. 1.7e-58;

Matches 141; Conservative 54; Mismatches 111; Indels 16; Gaps 3;

QY 9 RPKIAMVSGMIGTMAFLCSLRELGDVVLFDVVPNMPMGKAMDISHNSSVVDGTIVYG 68

DB 2 RKKISIIAGAFVGSITTAHMLAAKELGDIVLLDFVEGVPOGKALDLYEASPIEGFDRVTVG 61

QY 69 SNSYECLKADVVITAGITKIPGSKDKWSRMDLLPVNIKIMREVGAIAKSYCPNAPVI 128

DB 62 TNNYADTANSDIVVTSGAPRKP-----MSREDLIKVNADITRACISQAAPLSPNAVII 116

QY 129 NITNPLDVMAALQESSGLPHHRICGMAGMLDSSRFRMIADKLEVSPRDVGQVIGVHG 188

DB 117 MYNNPLDAMTYLAAEVSFGPKERVIGQAGVLDAAARYRTFIAMEAGVSVDVQAMLMGGHG 176

QY 189 DHMVPLSRYATVNGIPLSEFVKKGWIKQEEVDIVQKTKVAGGEIVRLLGQGSAYYAPGA 248

DB 177 DEWVPLPRESTISGIPVSEF-----IAPDLAQIVERTKGGGEIVNLLKTSAYYAPAA 231

QY 249 SAIQMAESYLKDRKRVWCSCYLOGQYGVQNHVLPVPCVIGRGVKEKIELELTAQEROE 308

DB 232 ATAQMVAEAVLKDKRVMVAAYLTGQYGLNDIYFGVPVILGAGGVKEKILELPLNEEMAL 291

QY 309 LOGSIDVEKEMQKATAALDASK 330

DB 292 LNASAKAVR-----ATLDTLK 307

RESULT 15

US-10-369-493-13585

; Sequence 13585, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 13585

; LENGTH: 319

; TYPE: PRT

; ORGANISM: Thermoplasma volcanium

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(319)

; OTHER INFORMATION: unsure at all Xaa locations

US-10-369-493-13585

Query Match 39.7%; Score 669.5; DB 4; Length 319;

Best Local Similarity 42.9%; Pred. No. 3.1e-58;

Matches 137; Conservative 68; Mismatches 91; Indels 23; Gaps 5;

QY 9 RPKIAMVSGMIGTMAFLCSLRELGDVVLFDVVPNMPMGKAMDISHNSSVVD 61

DB 3 RKKISVIGAGNVGATVAQFLATKELGDVVLFDVVDGIGPEKALDIOEGAPHWGLDGVV- 61

QY 62 TGITVYGSNSYECLKADVVITAGITKIPGSKDKWSRMDLLPVNIKIMREVGAIAKSY 121

DB 62 -GFSTSDSSNKNMEGSDVIVVTAGMARKPG-----MSREDLFDKNVEIADVSKNIKY 115

QY 122 CPNAPVINITNPLDVMAALQESSGLPHHRICGMAGMLDSSRFRMIADKLEVSPRDVG 181

```

Db 116 SPDSIIVVSNPADIMAYALQKISGVDQRIWGLGSLDSSRPTFLAKELDVSVEDVNA 175
Qy 182 MVIGVHGDHVPPLSRVATVNGIPLSEFVKKGWIKOBEVDDIVOKTVAGGEIVRLGQGS 241
Db 176 FVIGGHGDDMVFFIRYSSVAGIPIEKLLPK-----EKIDAIVKRTRFGGGEIVNYLKAGS 230
Qy 242 AYYAPGASAIOMAESYLKDKRVMVSCYLOGQ-----YGVQNHVYLGVPVIGRGVEKII 297
Db 231 AYYAPGISITAMWESVIKDKRVPICAAITGKHAHYGINNKFIGVPIKIGRGVEEY 290
Qy 298 ELELTAQEROELQGSIDEV 316
Db 291 DIDFPEELELWKKSVASV 309

```

Search completed: March 2, 2006, 19:45:39
Job time : 68 secs

This Page Blank (uspio)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 19:40:23 ; Search time 28 Seconds
(without alignments)
974.392 Million cell updates/sec

Title: US-09-390-846-2
Perfect score: 1688
Sequence: 1 MAVFEKTRPKIAVGSWMI.....GSIDEVEMQKAIKALDASK 330

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1688	100.0	330	2	US-08-676-882-2
2	828	49.1	315	4	PCT-US94-03796-2
3	706	41.8	329	1	US-08-270-0138-2
4	706	41.8	329	1	US-08-838-418-2
5	694	41.1	289	2	US-09-902-540-13486
6	641	38.0	325	2	US-09-134-001C-5533
7	464	27.5	324	2	US-09-107-532A-6486
8	426	25.2	318	2	US-09-134-000C-4417
9	412	24.4	317	1	US-08-748-068-3
10	412	24.4	333	1	US-08-748-068-1
11	406	24.1	354	2	US-09-949-016-8002
12	405	24.0	351	2	US-09-949-016-11252
13	405	24.0	351	2	US-09-949-016-11253
14	401.5	23.8	322	2	US-09-710-279-1412
15	401.5	23.8	322	2	US-09-710-279-2384
16	401.5	23.8	330	2	US-09-134-001A-4328
17	398	23.6	316	2	US-09-830-230A-561
18	395.5	23.4	307	2	US-09-107-532A-4274
19	391	23.2	333	1	US-08-869-506-2
20	391	23.2	333	1	US-09-128-967-2
21	390	23.1	333	1	US-08-869-506-3
22	390	23.1	333	2	US-09-128-967-3
23	389	23.0	331	2	US-09-711-681-4
24	389	23.0	331	2	US-10-274-266-4
25	388	23.0	331	2	US-09-107-433-4752
26	388	23.0	322	2	US-09-583-110-4591
27	377.5	22.4	327	1	US-08-748-068-2

28	371	22.0	295	2	US-09-830-230A-562	Sequence 562, App
29	371	22.0	381	2	US-09-711-681-2	Sequence 2, Appli
30	371	22.0	381	2	US-10-274-266-2	Sequence 2, Appli
31	363	21.5	320	2	US-09-535-381-2	Sequence 2, Appli
32	338	20.0	317	2	US-09-489-039A-9794	Sequence 9794, Ap
33	285.5	16.9	205	2	US-09-949-016-9136	Sequence 9136, Ap
34	277	16.4	339	2	US-09-543-681A-5827	Sequence 5827, Ap
35	273.5	16.2	304	2	US-09-634-238-238	Sequence 238, App
36	246.5	14.6	338	2	US-08-922-957-3	Sequence 3, Appli
37	240	14.2	338	2	US-08-922-957-1	Sequence 1, Appli
38	236.5	14.0	358	2	US-09-248-796A-17253	Sequence 17253, A
39	231.5	13.7	298	2	US-08-922-957-4	Sequence 4, Appli
40	225.5	13.4	262	2	US-09-949-016-9135	Sequence 9135, Ap
41	225.5	13.4	262	2	US-09-949-016-11622	Sequence 11622, A
42	224	13.3	333	2	US-09-248-796A-17254	Sequence 17254, A
43	217	12.9	343	2	US-09-248-796A-17252	Sequence 17252, A
44	212.5	12.6	322	2	US-09-565-501A-112	Sequence 112, App
45	212.5	12.6	322	2	US-09-639-206A-112	Sequence 112, App

ALIGNMENTS

RESULT 1
US-08-676-882-2
; Sequence 2, Application US/08676882
; Patent No. 6100241
; GENERAL INFORMATION:
; APPLICANT: Kok, Jacobus Johannes
; APPLICANT: van den Boogaart, Paul
; APPLICANT: Vermeulen, Arnoldus Nicolaas
; TITLE OF INVENTION: Coccidiosis poultry vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 6100241el Patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/676,882
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 258-5200
; TELEFAX: (301) 977-0847
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 330 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-676-882-2

Query Match	100.0%	Score 1688	DB 2	Length 330	
Best Local Similarity	100.0%	Pred. No. 4e-177	Indels 0	Gaps 0	
Matches 330	Conservative 0	Mismatches 0			
QY	1	MAVFEKTRPKIAVGSWMI	IGGTMAFLCSIR	ELGDLVFLFDVVVPM	PMGKAMDISHNSVV 60
DB	1	MAVFEKTRPKIAVGSWMI	IGGTMAFLCSIR	ELGDLVFLFDVVVPM	PMGKAMDISHNSVV 60
QY	61	DTGITVYGSNECLKAD	VITAGITKIPGKS	DKWSRMDLLP	VNIKIMREVGAAIKS 120

Db 61 DTGITVYGSNYSYCLKGADVVIIITAGITKIPGSKDKEWSDMLLPVNIKIMREVGAAIKS 120
QY 121 YCPNFAVINITNPLDVMVAALQESSGLPHHRIICGMAGMLDSSRFRMIADKLEVSRRDVQ 180
Db 121 YCPNFAVINITNPLDVMVAALQESSGLPHHRIICGMAGMLDSSRFRMIADKLEVSRRDVQ 180
QY 181 GMWIGHGDMVPLSRVATVNGIPLSEFVKKGWIKQEEVDIVQTKVAGGEIVRLIGQG 240
Db 181 GMWIGHGDMVPLSRVATVNGIPLSEFVKKGWIKQEEVDIVQTKVAGGEIVRLIGQG 240
QY 241 SAYYAPGASAIQMAESYLKDKRVMVCSYQLOGYGVQNHVYLVPCVIGRGVGEKIEIE 300
Db 241 SAYYAPGASAIQMAESYLKDKRVMVCSYQLOGYGVQNHVYLVPCVIGRGVGEKIEIE 300
QY 301 LTAQEROLOGSIDEVKEMOKATAALDASK 330
Db 301 LTAQEROLOGSIDEVKEMOKATAALDASK 330
RESULT 2
PCT-US94-03796-2
; Sequence 2, Application PC/TUS9403796
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: GENE ENCODING THE LACTATE DEHYDROGENASE
; TITLE OF INVENTION: ENZYME OF PLASMODIUM FALCIPARUM
; NUMBER OF SEQUENCES: 15
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/03796
; FILING DATE: 06-APR-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/046,160
; FILING DATE: 12-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: DCI-066CPCC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 315 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-03796-2

Query Match 49.1%; Score 828; DB 4; Length 315;
Best Local Similarity 52.6%; Pred. No. 2,1e-82;
Matches 164; Conservative 51; Mismatches 95; Indels 2; Gaps 1;
QY 9 RPKIAMVSGMIGTMAFLCSLRELGDVVLFDVVPNMPMGKAMDISHNSSVDDTGTIVYG 68
Db 4 KAKIVLVSGMIGVMTLIVQKNLGDVVLFDVIVQNPMPHGKALDTSHTNMYASNCKVSG 63
QY 69 SNSVECLKGADVVIIITAGITKIPGSKDKEWSDMLLPVNIKIMREVGAAIKSCPNAFVI 128
Db 64 SNTYDDLAGADVIVTAVGTAFKAPGSKDKEWNRDLPVNNKIMIEIGHHKKKCPNAFII 123
QY 129 NITNPLDVMVAALQESSGLPHHRIICGMAGMLDSSRFRMIADKLEVSRRDVQGVGVHG 188
Db 124 VTNPDVDMVQLLHQHSGVPKNIIGLGGVLDTSRLKYYISQKLVNCPVDVNAHIVGAHG 193
QY 189 DHMVPLSRVATVNGIPLSEFVKKGWIKQEEVDIVQTKVAGGEIVRLIGQGSAYYAPGA 248
Db 184 NKWLLKRYITVGGIPLQEFINNKLIISDALEAIFDRTVNTVALEIVNL--HASPYYVAPAA 241

QY 249 SAIQMAESYLKDKRVMVCSYQLOGYGVQNHVYLVPCVIGRGVGEKIEIELELTAEROE 308
Db 242 AIIEMAESYLKDKLVKLVICSTLEGGYHSDIFGGTPVVLGVANGVEQVIELQLNSEKAK 301
QY 309 LOGSIDEVKEMQ 320
Db 302 FDEAIAETKRMK 313
RESULT 3
US-08-270-013B-2
; Sequence 2, Application US/08270013B
; Patent No. 5686294
; GENERAL INFORMATION:
; APPLICANT: Sogabe et al.
; TITLE OF INVENTION: PROTEIN HAVING HEAT-RESISTANT MALATE
; TITLE OF INVENTION: DEHYDROGENASE ACTIVITY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 61601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/270,013B
; FILING DATE: 01-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 164701/1993
; FILING DATE: 02-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Robert F.
; REGISTRATION NUMBER: 27555
; REFERENCE/DOCKET NUMBER: 62321
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5600
; TELEFAX: (312) 616-5700
; TELEX: (25)3533
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 329 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-270-013B-2

Query Match 41.8%; Score 706; DB 1; Length 329;
Best Local Similarity 45.3%; Pred. No. 6.2e-69;
Matches 136; Conservative 64; Mismatches 88; Indels 12; Gaps 3;
QY 9 RPKIAMVSGMIGTMAFLCSLRELGDVVLFDV--VPNMPMGKAMDISHNSSVDDTGTIV 66
Db 5 RKKISVIGAGTGAATTAFLAQKELGDVVLVDIPLQENPTKGKALDMLSEASPVLGFDANI 64
QY 67 YGSNVECLKGADVVIIITAGITKIPGSKDKEWSDMLLPVNIKIMREVGAAIKSCPNAF 126
Db 65 IGTSYDADTADSDIVVITAGIARKPG-----MSRDDLVTTNOKIMKQVTKVVVYKSPNCY 119
QY 127 VNITNPLDVMVAALQESSGLPHHRIICGMAGMLDSSRFRMIADKLEVSRRDVQGVGV 186
Db 120 IIVLTFVDVMTTYTFKESGFPKRVIGSGVLDTARFRTFVAEELNISVKDVTGVVGG 179
QY 187 HGDHVMPLSRVATVNGIPLSEFVKKGWIKQEEVDIVQTKVAGGEIVRLIGQGSAYYAP 246
Db 180 HGDDMVPLVRSYVAGIPLKLEIPK-----DRLDAIVERTKGGGEIVNLLNGSAYYAP 234

Qy 247 GASAIQMAESYLKDRKRWVSCVLCQGYGVQNHVLPVCPVIGRGVKEKIELELTAQER 306
 Db 235 AASLVEMVEAILKQDRILPAIAYLEGEYGVGYLGVPTILGNGIEKIELELTAQER 294

RESULT 4

US-08-838-418-2
 ; Sequence 2, Application US/08838418
 ; Patent No. 5744342
 ; GENERAL INFORMATION:
 ; APPLICANT: Sogabe et al.
 ; TITLE OF INVENTION: PROTEIN HAVING HEAT-RESISTANT MALATE
 ; TITLE OF INVENTION: DEHYDROGENASE ACTIVITY
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
 ; STREET: Two Prudential Plaza, Suite 4900
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60601-6780.
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/838,418
 ; FILING DATE: 17-MAR-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 164701/1993
 ; FILING DATE: 02-JUL-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hoover, Allen E.
 ; REGISTRATION NUMBER: 37354
 ; REFERENCE/DOCKET NUMBER: 78339
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (312) 616-5600
 ; TELEFAX: (312) 616-5700
 ; TELEX: (25)3533
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 329 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-838-418-2

Query Match 41.8%; Score 706; DB 1; Length 329;
 Best Local Similarity 45.3%; Pred. No. 6.2e-69;
 Matches 136; Conservative 64; Mismatches 88; Indels 12; Gaps 3;
 Qy 9 RPKIAMVSGMIGTMAFLCSLRELGDVVLFDV--VPNMPGKAMDISHNSVVDTGITV 66
 Db 5 RKKISVIGAGTGAATFLAQKELGDVVLFDIPQLENPTGKALDMLASPLVGFDAI 64
 Qy 67 YGNSYSECLGADVITITAGITKIPGSKDKWRSMDLLPVNIKIMREVGAIAIKSYCPNAF 126
 Db 65 IGISDYADTADSDIVITAGIARKPG-----MSRDDLVTNNQIMKQVTEVKYISNCY 119
 Qy 127 VINITNPLDVMAALQESSGLPHHRICGMAGMLDSSRFRMIADKLEVSPRDVGMYIGV 186
 Db 120 IIVLTNPVDAMTVTFKESGPPKRVIGQSGVLDTRFRFTFAEELNISVKDVTGFLGG 179
 Qy 187 HGDHMYPLSYATVNGIPLSEFVKGWIKQEEVDIVQKTKVAGGEIVRLIGOGSAYAP 246
 Db 180 HGDDMVPLVRSYAGGIPLEKLPK-----DRLDAIVERTKGGGEIVNLLGNSAYAP 234

Qy 247 GASAIQMAESYLKDRKRWVSCVLCQGYGVQNHVLPVCPVIGRGVKEKIELELTAQER 306
 Db 235 AASLVEMVEAILKQDRILPAIAYLEGEYGVGYLGVPTILGNGIEKIELELTAQER 294

RESULT 5

US-09-902-540-13486
 ; Sequence 13486, Application US/09902540
 ; Patent No. 6833447
 ; GENERAL INFORMATION:
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Wiegand, Roger C.
 ; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
 ; FILE REFERENCE: 38-10(15849)B
 ; CURRENT APPLICATION NUMBER: US/09/902,540
 ; CURRENT FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: 60/217,883
 ; PRIOR FILING DATE: 2000-07-10
 ; NUMBER OF SEQ ID NOS: 16825
 ; SEQ ID NO 13486
 ; LENGTH: 289
 ; TYPE: PRT
 ; ORGANISM: Myxococcus xanthus
 ; US-09-902-540-13486

Query Match 41.1%; Score 694; DB 2; Length 289;
 Best Local Similarity 47.0%; Pred. No. 1.1e-67;
 Matches 140; Conservative 59; Mismatches 87; Indels 12; Gaps 3;
 Qy 6 KNTRPKIAMVSGMIGTMAFLCSLRELGDVVLFDV--VPNMPGKAMDISHNSVVDTGT 63
 Db 2 QNGKKIGLIGGGQIGGNLALLAVQKSLGDVLYDIPAAEGLVKALKALDINQLAADVGD 61
 Qy 64 ITVYGSNVECLGADVITITAGITKIPGSKDKWRSMDLLPVNIKIMREVGAIAIKSYCP 123
 Db 62 CRVKGTTDKVDAGSDVIIITAGMPRKPG-----MSREDLLEINLKIMTMDVAGNIKQHP 116
 Qy 124 NAFVINITNPLDVMAALQESSGLPHHRICGMAGMLDSSRFRMIADKLEVSPRDVGMY 183
 Db 117 NAFVINVAPLDANVFAHKKIAGLPKHMVAGMAGVLDTSRKFVABALGCSIRDVEALV 176
 Qy 184 IGVDHMYPLSYATVNGIPLSEFVKGWIKQEEVDIVQKTKVAGGEIVRLIGOGSAY 243
 Db 177 LGCHGDDMVPLVHRITVGGVPLTELAK-----DKLDAIKRTREGGPELVGLYKTSAY 231
 Qy 244 YAPGASAIQMAESYLKDRKRWVSCVLCQGYGVQNHVLPVCPVIGRGVKEKIELEL 301
 Db 232 FGPAACAIAMAESFLQDRKRVLPAAALLEGQYINGVYFGVPVQIGAGGVKEKIHVEL 289

RESULT 6

US-09-134-001C-5533
 ; Sequence 5533, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134,001C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 5533
 ; LENGTH: 325
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-5533

Query Match 38.0%; Score 641; DB 2; Length 325;
Best Local Similarity 41.6%; Pred. No. 8.8e-62;
Matches 132; Conservative 71; Mismatches 102; Indels 12; Gaps 4;

QY 9 RPKIAMVGSIGMTMAFLCSLRELGDVLFDD--VVPNMPMGKAMDISHNSSVVDGTGV 66
DB 16 RKISIIAGHTGGTLAFILAQKELGDVILIERQSEGMAKXALDILESQPTWGDTSV 75

QY 67 YGNSYSECLKGADVVIITAGITKIPGSKDSEWSDMLLPVNIKIMREVGAAIKSYCPNAP 126
DB 76 HGSVNIEDIKSDIVVMTAG---IPKSG--MTRRELVTQNEQIVRETALQIATYAPHSI 130

QY 127 VINITNPLDVMVAALQESSGLPHHRICMGAMLDSSRRFRMIADKLEVSPRDVGWIGV 186
DB 131 IIVLTNPDVDMVMTAFKASGFPKRIITGSGGILDAARYRTFIAQELNVSKVNGFVLGG 190

QY 187 HGDHVPPLSRVATVNGIPLSEFVKKGWIKOEVEDDVIQKTKVAGGEIVRLGGSAIYAP 246
DB 191 HGDMLPLNNTHINGIPV-----KHLISEKIDQIVERTKGAELVALLGGSAIYAP 245

QY 247 GASAIQMAESYLDKRRKVMVCSYQYQVQNHLYGVPVIGRGVKEKIELELTAQER 306
DB 246 ATAIVETIDAIFNDRKELPLSIAYLEGEYGCSDICFGVPTIIGVQGIKIEIIVDMNDEY 305

QY 307 QELQSGIDEVKEKQAI 323
DB 306 QQLQSAQAVSEVKNLSL 322

RESULT 7

US-09-107-532A-6486
; Sequence 6486, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneka
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6486:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 324 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...324
SEQUENCE DESCRIPTION: SEQ ID NO: 6486:
US-09-107-532A-6486

Query Match 27.5%; Score 464; DB 2; Length 324;
Best Local Similarity 32.8%; Pred. No. 2.8e-42;
Matches 106; Conservative 72; Mismatches 135; Indels 10; Gaps 5;

QY 3 VFEKTRPKIAMVGSIGMTMAFLCSLREL--GDVVLPDVPNMPMGKAMDISHNSSVVD 61
DB 9 LFMKKTSRKVVIVGTGFGVGTSIAYAMINQGISNELVLIDVNOKEKAGEALDLDGHWAGD 68

QY 62 TGITVYGSNSYSECLKGADVVIITAGITKIPGSKDSEWSDMLLPVNIKIMREVGAAIKSY 121
DB 69 ENVAVW--SGGYEECKDADIVITAGINQKPGQ-----SRDLVKTNASIMRQIVKEIMGS 122

QY 122 CNAFVINTNPLDVMVAALQESSGLPHHRICMGAMLDSSRRFRMIADKLEVSPRDVG 181
DB 123 GFDGIIVVASNPVDILTYIAMNESGLPTSRVIGTGTTLTTRFRKEIALKLKVDPRSVHG 182

QY 182 MWIGVHGDHVPPLSRVATVNGIPLSEFVKKGW--IKOEVEDDVIQKTKVAGGEIVRLGG 240
DB 183 YILGHEGDSEVAANSHTTVGGRPVFEIVKDHRIAQDELVDIADKVRNAAVELID--RKK 240

QY 241 SAYYAPGASAIQMAESYLDKRRKVMVCSYQYQVQNHLYGVPVIGRGVKEKIELE 300
DB 241 ATYYGIGMSTARIVKAILNNEQAVLPVSAYLTGEYDEKDFITGVPSIVDENGVRREVLS 300

QY 301 LTAQERQELQSGIDEVKEKQAI 323
DB 301 INEEKAMFSEKTSALREVINTV 323

RESULT 8

US-09-134-000C-4417
; Sequence 4417, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4417
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4417

Query Match 25.2%; Score 426; DB 2; Length 318;
Best Local Similarity 30.3%; Pred. No. 4.1e-38;
Matches 99; Conservative 80; Mismatches 130; Indels 18; Gaps 7;

QY 1 MAVFEKNTTRPKIAMVGSIGMTMAFLCSLRELG---DVLVFDVVPNMPMGKAMDISHNS 57
DB 2 MKVFNRK---KVAIIGTGFGVGTSIAY--SMINQGIANELILIDIDKAKSEGETIDLDGV 55

QY 58 SVVDGTGITVYGSNSYSECLKGADVVIITAGITKIPGSKDSEWSDMLLPVNIKIMREVGAA 117
DB 56 SWGQENNVNVMAGDYQNC--QNADIVIVITAGANQKPGQ-----SRDLVLSINAEIMKTIVNN 109

QY 118 IKSYCPNAPFVINITNPLDVMVAALQESSGLPHHRICMGAMLDSSRRFRMIADKLEVS 177

Db 110 IMKSGFDILVIASNPVDVLTYYVAQASGLPVSTIIIGTGTTLDTTRFKELSQRLAIDPR 169
Qy 178 DVQGMVIGVGHDMVPI-SRYATVNGIPLSEPV-KKGMIKOEVDVIVQKTKVAGGEIVRL 236
Db 170 NVHGYYIIGEGHSEVAVMSHTMTGKPILEIVDTTERTLSDDLPIISDKVKNTAYEIID- 228
Qy 237 LGQGSAYYAPGASAIOMAESYLDKDKRMVWCSCYLOGXGVONHYLGVPCVIGRGVEKI 296
Db 229 -RKQATYIGTGMSTARIVKAILNNEQAILPVSAIYLDGQYGOQDVFTGIPAVVNGQVTDI 287
Qy 297 IELETAQERQELQSGSIDEVEMQKAI 323
Db 288 IELNNAEKLFSKVTQLKQWASL 314

RESULT 9

US-08-748-068-3
Sequence 3, Application US/08748068
Patent No. 5770410
GENERAL INFORMATION:
APPLICANT: TITLE OF INVENTION: Chiral Synthesis
NUMBER OF SEQUENCES: 15
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/748,068
FILING DATE: 12-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/256,959
FILING DATE: 05-OCT-1994
APPLICATION NUMBER: GB 92 02033.8
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 92 04702.6
FILING DATE: 04-MAR-1992
APPLICATION NUMBER: GB 93/00204
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 317 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Modified-site
LOCATION: (15-16)
OTHER INFORMATION: /note= "- numbering discontinuity"
FEATURE:
NAME/KEY: Modified-site
LOCATION: (28-29)
OTHER INFORMATION: /note= "- numbering discontinuity"
FEATURE:
NAME/KEY: Modified-site
LOCATION: (29-30)
OTHER INFORMATION: /note= "- numbering discontinuity"
FEATURE:
NAME/KEY: Modified-site
LOCATION: (30-31)
OTHER INFORMATION: /note= "- numbering discontinuity"
FEATURE:
NAME/KEY: Modified-site
LOCATION: (40-41)
OTHER INFORMATION: /note= "- numbering discontinuity"

FEATURE:
NAME/KEY: Modified-site
LOCATION: (69-70)
OTHER INFORMATION: /note= "- numbering discontinuity"
FEATURE:
NAME/KEY: Modified-site
LOCATION: (73-74)
OTHER INFORMATION: /note= "- numbering discontinuity"
FEATURE:
NAME/KEY: Modified-site
LOCATION: (83-84)
OTHER INFORMATION: /note= "- numbering discontinuity"
FEATURE:
NAME/KEY: Modified-site
LOCATION: (115-116)
OTHER INFORMATION: /note= "- numbering discontinuity"
FEATURE:
NAME/KEY: Modified-site
LOCATION: (138-139)
OTHER INFORMATION: /note= "---- numbering discontinuity"
FEATURE:
NAME/KEY: Modified-site
LOCATION: (176-177)
OTHER INFORMATION: /note= "- numbering discontinuity"
FEATURE:
NAME/KEY: Modified-site
LOCATION: (181-182)
OTHER INFORMATION: /note= "- numbering discontinuity"
FEATURE:
NAME/KEY: Modified-site
LOCATION: (196-197)
OTHER INFORMATION: /note= "- numbering discontinuity"
FEATURE:
NAME/KEY: Modified-site
LOCATION: (211-212)
OTHER INFORMATION: /note= "- numbering discontinuity"
FEATURE:
NAME/KEY: Modified-site
LOCATION: (230-231)
OTHER INFORMATION: /note= "- numbering discontinuity"
FEATURE:
NAME/KEY: Modified-site
LOCATION: (234-235)
OTHER INFORMATION: /note= "- numbering discontinuity"
FEATURE:
NAME/KEY: Modified-site
LOCATION: (251-252)
OTHER INFORMATION: /note= "--- numbering discontinuity"
FEATURE:
NAME/KEY: Modified-site
LOCATION: (264-265)
OTHER INFORMATION: /note= "---- numbering discontinuity"
FEATURE:
NAME/KEY: Modified-site
LOCATION: (286-287)
OTHER INFORMATION: /note= "- numbering discontinuity"
US-08-748-068-3

Query Match 24.4%; Score 412; DB 1; Length 317;
Best Local Similarity 28.1%; Pred. No. 1.4e-36;
Matches 89; Conservative 79; Mismatches 139; Indels 10; Gaps 5;

Qy 5 EKNTRPKIAMVSGMIGGTWAFCLCSREIGD-VVLFDDVPMMPMGKAMDISHNSSVVDTC 63
Db 2 KNGGARVVVIGAGFVGASYVFALMNQGIADIEVLIDANESKAIGDAMDFNHGKVFAPKP 61
Qy 64 ITVYGSNSYECLKGADVLIITIGITIKIPGSKDSEWRMDLLPWIKIMREVGAAIKSYCP 123
Db 62 VDIWHDYDDC-RDADLVVICAGANQKPGE-----TRDLVKNKIAIFRSIVESVNASGF 115
Qy 124 NAFVINTNPLDVMVAALQESSGLPHHRICGMAGMLDSSRFRMIADKLEVSPRDVGWV 183


```
QY 178 DVQGMVIGVGHDMVPLSRVATVNGIPLSEF-----VKKGWIKOEVEDDDIVQKTKVA 229
Db 184 SCHGWLIGERGDSSVAVSGVNVAGVSLQQLNPMGTNDSENW---KEVHKMVVESAY- 239
QY 230 GGEIVRLGOGSAYYAPGASAIQMAESYLKDRKRVVVCSCYLOQYGVQNH-YLGVPCVI 288
Db 240 --EVIKL--KGYTNWAGLSVADLIESMLKLSRIHPVSTWVGMYGIENEVFLSLPCL 295
QY 289 GGRGVEKIELELTAQERQELQSGIDEVKEMQKAIAL 326
Db 296 NARGLTSVINQKLDKDEVAQLKNSADTLWGIQDKLDL 333

RESULT 11
US-09-949-016-8002
; Sequence 8002, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8002
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8002

Query Match 24.1%; Score 406; DB 2; Length 354;
Best Local Similarity 29.3%; Pred. No. 7.8e-36;
Matches 99; Conservative 75; Mismatches 130; Indels 34; Gaps 10;

QY 5 EKTRP--KIAMVSGMIGTMAFLCSLRELG-----DVVLFVDPVPMMPKAMDISHNS 57
Db 35 EEAATVPNNKITVVGQGVG---MACAISILGKSLADELALVDVLDKLGEMMDLQHS 90
QY 58 SVVDGTITVYGSNYECLKADVVITAGITKIPGSKDEWSRMDLLPVNIKIMRVGAA 117
Db 91 LFLQTP-KIVADKDYSVTANSKIVVVTAGVQOEGE-----SRNLVQRNVNVFKFIPO 144
QY 118 IKSCYPNAPVINITNPLDVMVAALQESSGLPHRIGCMAGMLDSSRRFRMIADKLEVS 177
Db 145 IVKSPDCIIIVSNPVDILTIVTWKLSGLPKHRVIGSGCNCNLSARFYLMAEKLGHPS 204
QY 178 DVQGMVIGVGHDMVPLSRVATVNGIPLSEF-----VKKGWIKOEVEDDDIVQKTKVA 229
Db 205 SCHGWLIGERGDSSVAVSGVNVAGVSLQQLNPMGTNDSENW---KEVHKMVVESAY- 260
QY 230 GGEIVRLGOGSAYYAPGASAIQMAESYLKDRKRVVVCSCYLOQYGVQNH-YLGVPCVI 288
Db 261 --EVIKL--KGYTNWAGLSVADLIESMLKLSRIHPVSTWVGMYGIENEVFLSLPCL 316
QY 289 GGRGVEKIELELTAQERQELQSGIDEVKEMQKAIAL 326
Db 317 NARGLTSVINQKLDKDEVAQLKNSADTLWGIQDKLDL 354

RESULT 12
US-09-949-016-11252
; Sequence 11252, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11252
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11252

Query Match 24.0%; Score 405; DB 2; Length 351;
Best Local Similarity 30.9%; Pred. No. 1e-35;
Matches 101; Conservative 72; Mismatches 130; Indels 24; Gaps 7;

QY 5 EKTRPKIAMVSGMIGTMAFLCSLRELCD-VVLFVDPVPMMPKAMDISHNSVVDTG 63
Db 35 DENSQCKITIVGTGAVGMACAISILKDLADELALVDVLDKLGEMMDLQHSLEFFST 94
QY 64 ITVYGSNYECLKADVVITAGITKIPGSKDEWSRMDLLPVNIKIMRVGAAIKSYCP 123
Db 95 KITSGKD-YVSANSRIVIVTAGARQOEGE-----TLALVQRNVNVAIKSIIPIVHYS 148
QY 124 NAFVINITNPLDVMVAALQESSGLPHRIGCMAGMLDSSRRFRMIADKLEVS 183
Db 149 DCKILVSNPVDILTIVTWKLSGLPKHRVIGSGCNCNLSARFYLIGEKLVHPTSCHGI 208
QY 184 IGVGHDMVPLSRVATVNGIPL-----SEFVKGWIKOEVEDDDIVQKTKVAGGIVR 235
Db 209 IGEHGDSSVPLMSGVNVAGVALKTLDPKLTGSDSDEKHW-----KNHKQVIOSEYBIK 262
QY 236 ILGQGSAYYAPGASAIQMAESYLKDRKRVVVCSCYLOQYGVQNH-YLGVPCVIGGRGVE 294
Db 263 LKGYTS--WAGLSVMDLVGSLKLRVRHPVSTWVGKLYGKBEFLSLPCLGRNGVS 320
QY 295 KIIELELTAQERQELQSGIDEVKEMQK 321
Db 321 DVVKINLSEERALFKKSAETLWNIQK 347

RESULT 13
US-09-949-016-11253
; Sequence 11253, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11253
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11253
```

```
Query Match      24.0%; Score 405; DB 2; Length 351;
Best Local Similarity 30.9%; Pred. No. 1e-35;
Matches 101; Conservative 72; Mismatches 130; Indels 24; Gaps 7;

QY 5 EKTRPKIAMVSGMTGGTMAFLCSURELGD-VVLFDDVVPNNPMGKAMDISHNSSVVDGTG 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 35 DENSQCKITIVGTGAVGMAICAILLKDLADELALVDVALDKLKGSMMDLQHGSLFFSTS 94
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 64 ITVYGSNECLGADVVIITAGITKIPGSKDEWSRMDLLPVNIKIMREVGNAIKSYCP 123
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 95 KITSGKD-YVSANSRIVITAGAQOGE-----TRLALVQRNVAIMKSIIPAIVHYSR 148
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 124 NAFVINITPLDVWVAALQESSGLPHHRICGMAGMLDSSRFRMIADKLEVSPRDQGMV 183
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 149 DCKILVSNPVDILTIVVKISGLPVRVIGSCNLD SARFYLIGELGVHTSCHKWI 208
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 184 IGVGHDMVPLSRATVNGIPL-----SEFVKKGWIKQEEVDDIVQKTKVAGGEIVR 235
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 209 IGEHGDSSVPLWGSVNVAGVALKTLDPKLGTDSDKEHW-----KNHKKVQISAVEIIK 262
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 236 LLQGSAYAPGASAIQMAESYLKDRKRVWVCYLGQYGVQNH-YLGVPCVIGGRGVE 294
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 263 LKGYTS--WAIGLSVMDLVGSLKNLRVHPVSTMVKGLYGIKEELFSLPCVLGRNGVS 320
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 295 KIIELELTAQERQELQSDIVKEMOK 321
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 321 DVVKINLNSEELFKKSAETLWNIOK 347
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
US-09-710-279-1412
; Sequence 1412, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1412
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-1412

Query Match      23.8%; Score 401.5; DB 2; Length 322;
Best Local Similarity 29.3%; Pred. No. 2.1e-35;
Matches 93; Conservative 75; Mismatches 140; Indels 9; Gaps 4;

QY 11 KIAMVSGMTGGTMAFLCSURELGD-VVLFDDVVPNNPMGKAMDISHNSSVVDGTGITYGS 69
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 13 KVVLVGDSGSSYAFAMVTQGIADFEFVIDIAKDKVEADVVDKLNHGALYSSSPVTV-KA 71
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 70 NSYECLKGADVVIITAGITKIPGSKDEWSRMDLLPVNIKIMREVGNAIKSYCPNAPVIN 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 72 GEYEDCKDADLVITAGAPQKPE-----TRLQLVEKNTKIMKSIIVTSVMDSGDFGFFLI 126
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 130 ITNPLDVWVAALQESSGLPHHRICGMAGMLDSSRFRMIADKLEVSPRDQGMVIGVHGD 189
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 127 AANPVDILTRYVKEVTGLPAERVIGSGTVLDSARFYLISKELGVTSVSSVHAIIGEHGD 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 190 HMVPLSRVATVNGIPLSEFVKKGWIKQEEVDDIVQKTKVAGGEIVRLLQGSAYIAPGAS 249
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 187 SELAVWSQANVGISVYDTLKEETGSDAKANEIYINTRDAAYDIIQ--AKGSTYYGIALA 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 250 AIQMAESYLKDRKRVWVCYLGQYGVQNH-YLGVPCVIGGRGVEKIIIELELTAQERQEL 309
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 245 LLRISKALLNNENSILTVSSQLNGQYGFNDVYLGTLTINQNGAVKIYETPLNDNELQLL 304
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 310 QGSIDEVEMQKAI AAL 326
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 305 EKS VKTLEDTYDSIKHL 321
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Search completed: March 2, 2006, 19:41:16
Job time : 29 secs
```


GenCore version 5.1.7

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 19:38:02 ; Search time 109 Seconds
(without alignments)

2136.003 Million cell updates/sec

Title: us-09-390-846-2

Perfect score: 1688

Sequence: 1 MAVFEKTRPKIAMVSGMI.....GSIDEVKEMKQIAALDASK 330

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1688	100.0	330	Q818U5_EIMAC	Q818U5 eimeria ace
2	1398	82.8	330	Q818U3_EIMMA	Q818U3 eimeria max
3	1266	75.0	331	Q818U4_EIMTE	Q818U4 eimeria ten
4	1072	63.5	326	LDH_TOXGO	Q27797 toxoplasma
5	1041.5	61.7	329	P90613_TOXGO	P90613 toxoplasma
6	905.5	53.6	316	Q6DQ12_TOXGO	Q6DQ12 toxoplasma
7	889	52.7	330	Q720X7_BABBO	Q720X7 babesia bov
8	886.5	52.5	320	MDH_RHLV	O33525 rhizobium l
9	881	52.2	316	Q4PRK9_PLAVI	O691m0 bartonella
10	877.5	52.0	320	MDH_BRHE	O8y7e7 brucella me
11	876.5	51.9	320	MDH_BRUME	Q7Cwk7 agrobacteri
12	873.5	51.7	320	MDH_AGR75	Q9ey76 rhizobium m
13	873.5	51.7	320	MDH_RHME	Q6fyd0 bartonella
14	870.5	51.6	320	MDH_BARQU	O57ax1 brucella ab
15	868.5	51.5	320	MDH_BRUAB	O89x59 bradyrhizob
16	868	51.4	322	MDH_BRAJA	O8fyf4 brucella su
17	863.5	51.2	320	MDH_BRUSU	P80458 rhodospseudo
18	862.5	51.1	322	MDH_RHOFA	Q4tp00 erythroba
19	862.5	51.1	324	Q4TP00_9SPHN	Q98ec4 rhizobium l
20	862	51.1	322	MDH_RHLO	Q84f58 methylbact
21	851.5	50.4	320	MDH_METEX	O51xel silicibacte
22	851.5	50.4	320	MDH_SILPO	Q9a2b1 caulobacter
23	850.5	50.4	320	MDH_CAUCR	Q6jh30 plasmodium
24	847	50.2	299	Q6JH30_PLAVI	Q7pbn3 rickettsia
25	846.5	50.1	314	Q7PBN3_RICSI	Q92ia0 rickettsia
26	837.5	49.6	314	MDH_RICCN	O7rhu8 plasmodium
27	836	49.5	316	Q7RHU8_PLAYO	O6ih31 plasmodium
28	835	49.5	299	Q6JH31_PLAMA	Q42783 plasmodium
29	833	49.3	316	Q42783_PLABE	Q7s197 plasmodium
30	833	49.3	322	Q7S197_PLABE	Q27743 plasmodium
31	831.5	49.3	316	LDH1_PLAPD	

RESULT 1

Q818U5_EIMAC PRELIMINARY; PRT; 330 AA.
 AC Q818U5;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Lactate dehydrogenase.
 GN Name=LDH;
 OS Eimeria acervulina.
 OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;
 OC Eimeria.
 OX NCBI_TaxID=5801;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Schaap D.C.;
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 [2]
 RN NUCLEOTIDE SEQUENCE.
 RA Kok H.J., van den Boogaart P., Vermeulen A.N., Schaap D.C.;
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 EMBL; AV143388; AAN38975.1; -; mRNA.
 DR HSP; Q27743; ICET.
 DR GO; GO:0004553; F:Hydrolase activity, hydrolyzing O-glycosyl . . . ; IEA.
 DR GO; GO:0004459; F:L-lactate dehydrogenase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR GO; GO:0006100; P:tricarboxylic acid cycle intermediate metab. . . ; IEA.
 DR InterPro; IPR001089; Glyco_hydro_4.
 DR InterPro; IPR001236; Ldh.
 DR InterPro; IPR001557; L_LDH_MDH.
 DR Pfam; PF02866; Ldh_1_C; 1.
 DR Pfam; PF00056; Ldh_1_N; 1.
 DR PIRSF; PIRSF000102; Lac dehydrog; 1.
 DR PRINTS; PD00086; LLDHGRNASE.
 DR ProDom; PD006892; Glyco_hydro_4; 1.
 KW Oxidoreductase.
 SQ SEQUENCE 330 AA; 35857 MW; 79500C3360E3A84C CRC64;

Query Match 100.0%; Score 1688; DB 2; Length 330;

Best Local Similarity 100.0%; Pred. No. 8.7e-120;

Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVFEKTRPKIAMVSGMIGGTMAFLCSLRELGDVVLPDVVPMMPMGKAMDLSHNSVV 60
 |||||
 DB 1 MAVFEKTRPKIAMVSGMIGGTMAFLCSLRELGDVVLPDVVPMMPMGKAMDLSHNSVV 60
 |||||
 QY 61 DTGITVYGSNSYECLGADVVIITAGITKIPGSKDSEWSEMDLLPVNIKIMREVGAAIKS 120
 |||||
 DB 61 DTGITVYGSNSYECLGADVVIITAGITKIPGSKDSEWSEMDLLPVNIKIMREVGAAIKS 120
 |||||
 QY 121 YCNPAFVINITNPLDVNVAALQESSGLPHHRICMGAGMLDSSRRFRRIADKLEVPDQV 180
 |||||

```
Db 121 YCPNFAVINITNPLDVMVAALQESSGLPHHRIICGMAGMLDSSRRPRMIADKLEVSPRDVQ 180
Qy 181 GMVGVGHGDMVPLSRVATVNGIPLSEFVKKGWIKQBEVDDIVQKTKVAGGEIVRLLGQ 240
Db 181 GMVGVGHGDMVPLSRVATVNGIPLSEFVKKGWIKQBEVDDIVQKTKVAGGEIVRLLGQ 240
Qy 241 SAYYAPGASATQMAESYLKDKRKVMVCSYVLOGQYGVQNHVYLGVPVCGVGGVEKIELE 300
Db 241 SAYYAPGASATQMAESYLKDKRKVMVCSYVLOGQYGVQNHVYLGVPVCGVGGVEKIELE 300
Qy 301 LTAQERQELQGSIDEVKEMQKAIALDASK 330
Db 301 LTAQERQELQGSIDEVKEMQKAIALDASK 330

RESULT 2
ID Q818U3_EIMMA PRELIMINARY; PRT; 330 AA.
AC Q818U3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Lactate dehydrogenase.
GN Name=Ldh;
OS Eimeria maxima.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;
OC Eimeria.
OX NCBI_TaxID=5804;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Schaap D.C.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Niesen R., Schaap D.C.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY143390; AAN38977.1; -, mRNA.
DR HSSP; Q27743; 1CET.
DR GO; GO:0004553; F:hydrolyase activity, hydrolyzing O-glycosyl . . . ; IEA.
DR GO; GO:0004459; F:L-lactate dehydrogenase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR GO; GO:0006100; P:tricarboxylic acid cycle intermediate metab. . . ; IEA.
DR InterPro; IPR001088; Glyco_hydro_4.
DR InterPro; IPR001236; ldh.
DR InterPro; IPR001557; L_LDH_MDH.
DR Pfam; PF02866; Ldh_1_C; 1.
DR Pfam; PF00056; Ldh_1_N; 1.
DR PIRSF; PIRSF00102; Lac dehydrog; 1.
DR PRINTS; PR00086; LDHDRGNASE.
DR ProDom; PD006892; Glyco_hydro_4; 1.
DR Oxidoreductase.
SQ SEQUENCE 330 AA; 35951 MW; 380CB5B7151B59E8 CRC64;

Query Match 82.8%; Score 1398; DB 2; Length 330;
Best Local Similarity 80.6%; Pred. No. 9e-98;
Matches 266; Conservative 29; Mismatches 35; Indels 0; Gaps 0;

Qy 1 MAVFEQKTRPKIAMVSGMIGTMAFLCSLRELGDVVLFDVVPNMPGKAMDISHNSVV 60
Db 1 MAVFEQKTRPKIALVSGMIGTMAFLCSLRELGDVVLFDVVPNMPGKAMDLSHNSVV 60

Qy 61 DTGITVYGSNSYECLTGADVVIIITAGITKIPGSKDKEWRMDLLPVNIKIMREVGAAIKS 120
Db 61 DNGITVYGSNSYECLTNADVVIIITAGITKIPGSKDKEWRMDLLPVNIKIMREVGAAIKK 120

Qy 121 YCPNFAVINITNPLDVMVAALQESSGLPHHRIICGMAGMLDSSRRPRMIADKLEVSPRDVQ 180
Db 121 YCPNFAVINITNPLDVMVAALQESSGLPHHRIICGMAGMLDSSRRPRMIADCLHVSHPDVQ 180

Qy 181 GMVGVGHGDMVPLSRVATVNGIPLSEFVKKGWIKQBEVDDIVQKTKVAGGEIVRLLGQ 240
Db 181 GMVGVGHGDMVPLSRVATVNGIPLSEFVKKGWIKQBEVDDIVQKTKVAGGEIVRLLGQ 240
```

```
Db 181 GMVGVGHGDMPLMRYITINGIPIQEFINKGLINKKEINNIYKTKQAGGDIVRLLGQ 240
Qy 241 SAYYAPGASATQMAESYLKDKRKVMVCSYVLOGQYGVQNHVYLGVPVCGVGGVEKIELE 300
Db 241 SAYYAPGTSAILMAESYLKDKRFLFVSSCYNGQYNNVNNHYLGVPVCIIGKGIEFIELD 300
Qy 301 LTAQERQELQGSIDEVKEMQKAIALDASK 330
Db 301 LNOEEKLLQGSIDEVLEMKQKAIALDAGK 330

RESULT 3
ID Q818U4_EIMTE PRELIMINARY; PRT; 331 AA.
AC Q818U4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Lactate dehydrogenase.
GN Name=Ldh;
OS Eimeria tenella.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;
OC Eimeria.
OX NCBI_TaxID=5802;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Schaap D.C.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Arts G., Kroezen H., Schaap D.C.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the LDH/MDH superfamily. LDH family.
DR EMBL; AY143389; AAN38976.1; -, mRNA.
DR HSSP; Q27743; 1CET.
DR GO; GO:0004459; F:L-lactate dehydrogenase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR GO; GO:0006100; P:tricarboxylic acid cycle intermediate metab. . . ; IEA.
DR InterPro; IPR001236; ldh.
DR InterPro; IPR001557; L_LDH_MDH.
DR Pfam; PF02866; Ldh_1_C; 1.
DR Pfam; PF00056; Ldh_1_N; 1.
DR PIRSF; PIRSF00102; Lac dehydrog; 1.
DR PRINTS; PR00086; LDHDRGNASE.
DR Oxidoreductase.
SQ SEQUENCE 331 AA; 34965 MW; BEF87B9F837AE469 CRC64;

Query Match 75.0%; Score 1266; DB 2; Length 331;
Best Local Similarity 71.4%; Pred. No. 9.4e-88;
Matches 235; Conservative 51; Mismatches 43; Indels 0; Gaps 0;

Qy 1 MAVFEKTRPKIAMVSGMIGTMAFLCSLRELGDVVLFDVVPNMPGKAMDISHNSVV 60
Db 1 MAVFEKVRPKIALVSGMIGTMAFLCSLRELGDVVLFDVVPNMPAGKALDCHTAAVA 60

Qy 61 DTGITVYGSNSYECLTGADVVIIITAGITKIPGSKDKEWRMDLLPVNIKIMREVGAAIKS 120
Db 61 DNGVRVQGANSYASLEGADVIIITAGITTKAAGSKDQESKRDLLPVNVTILREVGAIKQ 120

Qy 121 YCPNFAVINITNPLDVMVAALQESSGLPHHRIICGMAGMLDSSRRPRMIADKLEVSPRDVQ 180
Db 121 FCPHAFVINITNPLDVMVAALREAAGLPAARVCGMAGVLDSARFRLLADRLGVSPRDVQ 180

Qy 181 GMVGVGHGDMVPLSRVATVNGIPLSEFVKKGWIKQBEVDDIVQKTKVAGGEIVRLLGQ 240
Db 181 AMVLGVGHGDMVPLSRFATVNGVPLGELARQGWISEATREVERQTRAGGDIVRLLGQ 240

Qy 241 SAYYAPGASATQMAESYLKDKRKVMVCSYVLOGQYGVQNHVYLGVPVCGVGGVEKIELE 300
Db 241 SAYFAFGAAVAAAEAYLKDKRFVFCVSCYLEGPGVGRHCLGVPVCGVGGVERVIELP 300

Qy 301 LTAQERQELQGSIDEVKEMQKAIALDAS 329
```

Db 301 LDAREAQLQASIDEVREHMQLAADAA 329

RESULT 4

LDH_TOXGO

ID LDH_TOXGO STANDARD; PRT; 326 AA.

AC Q27797;

AC 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE L-lactate dehydrogenase (EC 1.1.1.27) (LDH).

OS Toxoplasma gondii.

OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Sarcocystidae;

OC Toxoplasma.

OX NCBI_TaxID=5811;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=ME49;

RC MEDLINE=96123406; PubMed=8577343; DOI=10.1016/0166-6851(95)00069-D;

RA "A bradyzoite stage-specifically expressed gene of Toxoplasma gondii encodes a polypeptide homologous to lactate dehydrogenase.";

RL Mol. Biochem. Parasitol. 73:291-294 (1995).

CC -1- CATALYTIC ACTIVITY: (S)-lactate + NAD(+) = pyruvate + NADH.

CC -1- PATHWAY: Anaerobic glycolysis; final step.

CC -1- SUBUNIT: Homotetramer (By similarity).

CC -1- SIMILARITY: Belongs to the LDH/MDH superfamily. LDH family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

DR EMBL; U23207; AAC46863.1; -, mRNA.

DR PDB; ISOV; X-ray; A/B=1-326.

DR SNR; Q27797; 3-325.

DR InterPro; IPR011304; L-LDH-NAD.

DR InterPro; IPR001557; L_LDH_MDH.

DR PANTHER; PTHR11540; Ldh; 1.

DR Pfam; PF02866; Ldh_1_C; 1.

DR Pfam; PF00056; Ldh_1_N; 1.

DR PIRSF; PIRSF00102; Lac mal DH; 1.

DR PRINTS; PR00086; LLDHGRGNASE.

DR PROSITE; PS00064; L_LDH; FALSE NEG.

KW 3D-structure; Glycolysis; NAD; Oxidoreductase.

FT NP_BIND 39 60

FT ACT_SITE 186 186 Proton acceptor (By similarity).

FT BINDING 99 99 Substrate (By similarity).

FT BINDING 131 131 NAD or substrate (By similarity).

FT BINDING 162 162 Substrate (By similarity).

SQ SEQUENCE 326 AA; 35307 MW; E692C95A81FC031E CRC64;

Query Match 63.5%; Score 1072; DB 1; Length 326;

Best Local Similarity 62.6%; Pred. No. 4.9e-73;

Matches 199; Conservative 56; Mismatches 63; Indels 0; Gaps 0;

QY 9 RPKIAMVSGMIGTGMATFLCSRLRELGDVVLFDVPPNMPMGKAMDISHNSVVDGTGITVYG 68

DB 8 RKKIAMIGSGMIGTGMATFLCSRLRELGDVVLFDVPPNMPMGKAMDISHNSVVDGTGITVYG 67

QY 69 SNSYECLKADVVIITAGITKIPKSKDEWSRMDLLPWNKIMREVGAAIKSYCPNAFVI 128

DB 68 ANQYEKIAGSDVVIITAGITKIPKSKDEWSRMDLLPWNKIMREVGAAIKSYCPNAFVI 127

QY 129 NITNPLDVMVAALQESSGLPHHRIICMGAGMLDSSRFMRMTADKLEVSPRDVQGMVGVGH 188

DB 128 VVTNPLDCMVKCFEASGLPKMVMCGMANVLDSAPFRFIADQLEISPRDIQATVIGTHG 187

QY 189 DHMVLPSRYATVNGIPLSEFVKKGWIKQEEVDDIVQKTKVAGGEIVRLLLGQGSAYAPGA 248

Db 301 LDAREAQLQASIDEVREHMQLAADAA 329

RESULT 5

P90613_TOXGO

ID P90613_TOXGO PRELIMINARY; PRT; 329 AA.

AC P90613;

AC 01-MAY-1997 (Tremblrel. 03, Created)

DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)

DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)

DE Lactate dehydrogenase (EC 1.1.1.27).

OS Toxoplasma gondii.

OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Sarcocystidae;

OC Toxoplasma.

OX NCBI_TaxID=5811;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=ME49;

RC MEDLINE=97169139; PubMed=9016946; DOI=10.1016/S0378-1119(96)00566-5;

RA Yang S., Parmley S.F.;

RT "Toxoplasma gondii expresses two distinct lactate dehydrogenase homologous genes during its life cycle in intermediate hosts.";

RL Gene 184:1-12(1997).

CC -1- SIMILARITY: Belongs to the LDH/MDH superfamily. LDH family.

DR EMBL; U35118; AAC47443.1; -, mRNA.

DR PDB; 1PZE; X-ray; A=1-329.

DR PDB; 1PZF; X-ray; A/B/C/D=1-329.

DR PDB; 1PZG; X-ray; A/B/C/D=1-329.

DR PDB; 1PZH; X-ray; A/B/C/D=1-329.

DR GO; GO:0004459; F-L-lactate dehydrogenase activity; IEA.

DR GO; GO:0016491; P-oxidoreductase activity; IEA.

DR GO; GO:0006096; P-glycolysis; IEA.

DR GO; GO:0006100; P-tricarboxylic acid cycle intermediate metab. .; IEA.

DR InterPro; IPR001236; Ldh.

DR InterPro; IPR001557; L_LDH_MDH.

DR Pfam; PF02866; Ldh_1_C; 1.

DR Pfam; PF00056; Ldh_1_N; 1.

DR PIRSF; PIRSF00102; Lac dehydrog; 1.

DR PRINTS; PR00086; LLDHGRGNASE.

KW Oxidoreductase.

SQ SEQUENCE 329 AA; 35548 MW; F0638785AC0E6BBB CRC64;

Query Match 61.7%; Score 1041.5; DB 2; Length 329;

Best Local Similarity 60.4%; Pred. No. 1e-70; 71; Indels 1; Gaps 1;

Matches 194; Conservative 55; Mismatches 71; Indels 1; Gaps 1;

QY 9 RPKIAMVSGMIGTGMATFLCSRLRELGDVVLFDVPPNMPMGKAMDISHNSVVDGTGITVYG 68

DB 9 RKKIAMIGSGMIGTGMATFLCSRLRELGDVVLFDVPPNMPMGKAMDISHNSVVDGTGITVYG 68

QY 69 SNSYE-CLKGADVVIITAGITKIPKSKDEWSRMDLLPWNKIMREVGAAIKSYCPNAFV 127

DB 69 EYSYEALATGADCVITVITAGITKIPKSKDEWSRMDLLPWNKIMREVGAAIKSYCPNAFV 128

QY 128 INITNPLDVMVAALQESSGLPHHRIICMGAGMLDSSRFMRMTADKLEVSPRDVQGMVGVGH 187

DB 129 IVTNPDLDCMVKCFEASGLPKMVMCGMANVLDSAPFRFIADQLEISPRDIQATVIGTHG 188

QY 188 GDMVPLSRATVNGIPLSEFVKKGWIKQEEVDDIVQKTKVAGGEIVRLLLGQGSAYAPG 247

DB 189 GDMVPLSRATVNGIPLSEFVKKGWIKQEEVDDIVQKTKVAGGEIVRLLLGQGSAYAPG 248

QY 248 ASAIOMASSYLKDKRVMVCSYLOGQYGVNHVLPVCVITGGRGVEKIELELTATQERQ 307

```

Db      249 ASAVAMATSFNLNDEKRVPCSVYCNGBVGLKDMFGLPAVIGGAGIERVIELELNEBEKK 308
Qy      308 ELQGSIDEVEMQKAIALDA 328
Db      309 QFQKSVDDVNMALNKAVAAQA 329

RESULT 6
Q6DQL2_TOXGO
ID      Q6DQL2_TOXGO PRELIMINARY;      PRT;      316 AA.
AC      Q6DQL2;
DT      25-OCT-2004 (TrEMBLrel. 28, Created)
DT      25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT      13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE      Malate dehydrogenase.
OS      Toxoplasma gondii.
OC      Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Sarcocystidae;
OC      Toxoplasma.
OX      NCBI_TaxID=5811;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=RH;
RA      Shen C., Zhan X., He A., Li Z.;
RL      Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=RH;
RX      PubMed=14598170; DOI=10.1007/s00436-003-0996-1;
RA      Chan M., Sim T.S.;
RT      "Functional characterization of an alternative (lactate dehydrogenase-
RT      like) malate dehydrogenase in Plasmodium falciparum.";
RL      Parasitol. Res. 92:43-47(2004).
RN      [3]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=RH;
RA      Shen C., Zhan X., Zheng X., Zheng B., He A., Li Z., Zheng H.;
RT      "Cloning and characterization of the malate dehydrogenase gene from
RT      Toxoplasma gondii.";
RL      Zhongguo Ren Shou Gong Huan Bing Za Zhi 21:0-0(2005).
DR      EMBL; AY650028; AAT67462.1; -; mRNA.
DR      EMBL; AY972525; AAX83280.1; -; Genomic DNA.
DR      GO; GO:0004459; F:L-lactate dehydrogenase activity; IEA.
DR      GO; GO:0030060; F:L-malate dehydrogenase activity; IEA.
DR      GO; GO:0016491; F:oxidoreductase activity; IEA.
DR      GO; GO:0006096; P:glycolysis; IEA.
DR      GO; GO:0006108; P:malate metabolism; IEA.
DR      InterPro; IPR001236; ldh.
DR      InterPro; IPR01557; L_LDH MDH.
DR      InterPro; IPR011275; MalateDH_bact.
DR      Pfam; PF02866; Ldh_1_C; 1.
DR      Pfam; PF00056; Ldh_1_N; 1.
DR      PIRSF; PIRSF00102; Lac dehydrog; 1.
DR      PRINTS; PRO0086; LLDHDRGNASE.
DR      TIGRFAMs; TIGR01763; MalateDH_bact; 1.
SQ      SEQUENCE 316 AA; 33777 MW; 10361A049B71D559 CRC64;

Query Match      53.6%; Score 905.5; DB 2; Length 316;
Best Local Similarity 54.9%; Pred. No. 2.le-60;
Matches 173; Conservative 57; Mismatches 80; Indels 5; Gaps 1;

Qy      8 TRPKIAMVSGMTGGTMAFLCSLRELGDVVLFDVVPNMPMGKAMDISHNSVVDGTGITYV 67
Db      2 SRRKIGLIGNICATLALLSVAVKELGDVNVFDVVDLPQKCLDLQLYLPISGVDVRF 61
Qy      68 GSNSEYCLKGADVVIITAGITKIPGSKDKEWSRMDLLPVNIIKIMREYGAIAIKSCYCPNAFV 127
Db      62 GSDNSYVLKQADVLIIVTAGVPRKFG-----MSRDDLLAINAKINGQVGEAIKQYCPNAFV 116
Qy      128 INTNPLDVMVAALQESSGLPHHRICMAGMLDSSRRFRMIADKLEYSPRDVGQMVJGVH 187
Db      117 ICITNPLDVMVYILREKCGLPHPKVCVMAGVLDLSARLRTFLSERLNVSVDDIHALVMG 176
Qy      188 GDHVPILSRVATVNGIPLSEFVKKGWIKQBEVDVIVQTKVAGGEIVRLIGQGSAYYAPG 247

Query Match      52.7%; Score 889; DB 2; Length 330;
Best Local Similarity 52.7%; Pred. No. 3.9e-59;
Matches 166; Conservative 62; Mismatches 87; Indels 0; Gaps 0;

Qy      9 RPKIAMVSGMTGGTMAFLCSLRELGDVVLFDVVPNMPMGKAMDISHNSVVDGTGITYV 68
Db      15 RNKISLIGSGNIGGVNAYLAQLKELGDVVLFDIAPKLGAAKALDIMHANAYDTQSNVIG 74
Qy      69 SNSYECLKGADVVIITAGITKIPGSKDKEWSRMDLLPVNIIKIMREYGAIAIKSCYCPNAFV 128
Db      75 TTSYEDIAGSDVCIITAGLAKLPNKSDDSDSDLDLVPAPNSKINFTIGENIKKYAPNAFV 134
Qy      129 NITNPLDVMVAALQESSGLPHHRICMAGMLDSSRRFRMIADKLEYSPRDVGQMVJGVH 188
Db      135 CITNPLDVMVKKLLKSTGTFPKNVVGMGGLDSSRMCHYIADKLRVNPVYVHSGCIGHG 194
Qy      189 DHMVPILSRVATVNGIPLSEFVKKGWIKQBEVDVIVQTKVAGGEIVRLIGQGSAYYAPGA 248
Db      195 DSMIPLTNHTVNGIPIQRIERGETIQAELDKIAERTIGSGMELVQLYNGSAFFAPAT 254
Qy      249 SATQMAESYLKORKRVNVCYLOGQYQVNHVLPVPCVIGRGVGEIKIIELELTAOERQ 308
Db      255 AALEMASAYLSDKRSVIVCSYLEGEGYHNDVLTGTPAIGANGIEKIITLKLSABEQAK 314
Qy      309 LOGSIDEVEMQKAI 323
Db      315 LDASVKEIRRLALV 329

```

```

RESULT 8
MDH_RHLV
ID MDH_RHLV STANDARD; PRT; 320 AA.
AC O33525;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Malate dehydrogenase (EC 1.1.1.37).
GN Name=mdh;
OS Rhizobium leguminosarum bv. viciae.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=387;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=3841;
RA Poole P.S., Allaway D., Smith M.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Catalyzes the reversible oxidation of malate to
CC oxaloacetate (By similarity).
CC -!- CATALYTIC ACTIVITY: (S)-malate + NAD(+) = oxaloacetate + NADH.
CC -!- SIMILARITY: Belongs to the LDH/MDH superfamily. MDH type 3 family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AJ002750; CAA05717.1; -; Genomic DNA.
CC HSSP; P80040; 1GUU.
CC HAMAP; MF_00487; -; 1.
CC InterPro; IPR001557; L_LDH_MDH.
CC InterPro; IPR001236; Ldh.
CC InterPro; IPR011275; MalateDH_bact.
CC InterPro; IPR001252; Mdh_AS.
CC PANTHER; PTHR11540; ldh; 1.
CC Pfam; PF02866; Ldh_1_C; 1.
CC Pfam; PF00056; Ldh_1_N; 1.
CC PIRSF; PIRSF00102; Lac_mal_DH; 1.
CC PRINTS; PR00086; LLDHGRNAS5.
CC TIGRfams; TIGR01763; MalateDH_bact; 1.
KW NAD; Oxidoreductase; Tricarboxylic acid cycle.
FT ACT_SITE 149 149 Proton relay (By similarity).
FT ACT_SITE 176 176 Proton relay (By similarity).
FT BINDING 83 83 Substrate carboxyl (By similarity).
FT BINDING 89 89 Substrate carboxyl (By similarity).
FT BINDING 152 152 Substrate carboxyl (By similarity).
SQ SEQUENCE 320 AA; 33590 MW; 2094407A42C2B1C8 CRC64;

Query Match 52.5%; Score 886.5; DB 1; Length 320;
Best Local Similarity 54.5%; Pred. No. 5.8e-59;
Matches 171; Conservative 59; Mismatches 79; Indels 5; Gaps 1;

QY 9 RPKIAMVSGMIGTMAFLCSRLGDLVDFVVPNNPMGKAMDISHNSSVVDGTGIVYG 68
DB 3 RNKIALIGSGMIGTMAFLCSRLGDLVDFVVPNNPMGKAMDISHNSSVVDGTGIVYG 62
QY 69 SNSYECLKGADVVITAGITKIPKSKDKEWSRMDLLFPVNIKIMREVGAAIKSYCPNFAVI 128
DB 63 ASDYSATGADVCIVTAGVARKPG-----NSRDLLGLNLKWEQVAGIKKYAPNFAVI 117
QY 129 NITNPLDMVVAALQESSGLPHRRICGMAGMLDSSRRFRMIADKLEVSPRDQGVVGVHG 188
DB 118 CITNPLDMVVAALQESSGLPHRRICGMAGMLDSSRRFRMIADKLEVSPRDQGVVGVHG 177
QY 189 DHMVPLSRVATVNGIPLSEFVKKGWIKQBEVDVIVOKTKVAGGEIVRLLGGSAYVAPGA 248
DB 178 DTWVPLARYSTVGIPITDLVTMGVWTKERLEEIIQTRDGGAEIVGLLTKGSAAYAPAA 237
QY 249 SAIQMAESYLKDRKRMVVCSCYLQGGYGVQNHVLPVPCVIGRGVKEKIELELTAQERQE 308

RESULT 9
MDH_BARHE
ID MDH_BARHE STANDARD; PRT; 320 AA.
AC Q6GIM0;
DT 25-OCT-2004 (Rel. 45, Created)
DB 238 SAIEMAESYLKDRKRMVVCSCYLQGGYGVQNHVLPVPCVIGRGVKEKIELELTAQERQE 297
QY 309 LQGSIDEVEMQKAI 322
DB 298 FDKSVGVAGLCEA 311
RESULT 9
Q4PRK9 PLAVI PRELIMINARY; PRT; 316 AA.
AC Q4PRK9;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE L-lactate dehydrogenase (EC 1.1.1.27).
GN Name=LDH;
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Belem;
RA PubMed-15218378; DOI=10.1023/B:BILE.0000032958.78158.10;
RA Turgut-Balik D., Akbulut E., Shoemark D.K., Celik V., Moreton K.M.,
RA Sessions R.B., Holbrook J.J., Brady R.L.;
RT "Cloning, sequence and expression of the lactate dehydrogenase gene
RT from the human malaria parasite, Plasmodium vivax.";
RL Biotechnol. Lett. 26:1051-1055(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Belem;
RA Turgut-Balik D., Akbulut E., Shoemark D.K., Celik V., Moreton K.M.,
RA Sessions R.B., Brady L.R.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ60151; AAY59419.1; -; Genomic DNA.
KW Oxidoreductase.
SQ SEQUENCE 316 AA; 34223 MW; 9DFA101BF9B9E291D CRC64;

Query Match 52.2%; Score 881; DB 2; Length 316;
Best Local Similarity 55.9%; Pred. No. 1.1e-58;
Matches 176; Conservative 48; Mismatches 89; Indels 2; Gaps 1;

QY 9 RPKIAMVSGMIGTMAFLCSRLGDLVDFVVPNNPMGKAMDISHNSSVVDGTGIVYG 68
DB 4 KPKIVLVGSGMIGTMAFLVQNLGDLVDFVVPNNPMGKALDTSNWNVWVNSCKVTG 63
QY 69 SNSYECLKGADVVITAGITKIPKSKDKEWSRMDLLFPVNIKIMREVGAAIKSYCPNFAVI 128
DB 64 SNSYDDLKGADVVITAGITKIPKSKDKEWSRMDLLFPVNIKIMREVGAAIKSYCPNFAVI 123
QY 129 NITNPLDMVVAALQESSGLPHRRICGMAGMLDSSRRFRMIADKLEVSPRDQGVVGVHG 188
DB 124 VTNPDVVMVQLLFHSGVPPKNIIGLVGLDTSRLKYYISOKLVNCPDVDNALVGAHG 183
QY 189 DHMVPLSRVATVNGIPLSEFVKKGWIKQBEVDVIVOKTKVAGGEIVRLLGGSAYVAPGA 248
DB 184 NKMVLLKRYITVGGIPLQEFINNKITDEVEGIFRTVNTALEIVNLL--ASPYVAPAA 241
QY 249 SAIQMAESYLKDRKRMVVCSCYLQGGYGVQNHVLPVPCVIGRGVKEKIELELTAQERQE 308
DB 242 AIEMAESYLKDIKKVLCVSTLLEGGYGHNSIFGGTPLVIGTGGVQVIEQLNAEETK 301
QY 309 LQGSIDEVEMQKAI 323
DB 302 FDEAVAEATKMKALI 316

RESULT 10
MDH_BARHE
ID MDH_BARHE STANDARD; PRT; 320 AA.
AC Q6GIM0;
DT 25-OCT-2004 (Rel. 45, Created)
```

```
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Malate dehydrogenase (EC 1.1.1.37)
GN Name=mdh; OrderedLocusNames=BH16570;
OS Bartonella henselae (Rochalimaea henselae).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bartonellaceae; Bartonella.
OX NCBI_TaxID=39323;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 49882 / Houston 1;
RX PubMed=15210978; DOI=10.1073/pnas.0305659101;
RA Alamarck U.C.M., Frank A.C., Karlberg E.O., Legault B.-A., Ardell D.H.,
RA Canbaeck B., Eriksson A.-S., Naeslund A.K., Handley S.A., Huvt M.,
RA La Scoia B., Holmberg M., Andersson S.G.E.;
RT "The house-borne human pathogen Bartonella quintana is a genomic
RT derivative of the zoonotic agent Bartonella henselae.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9716-9721(2004).
CC -|- FUNCTION: Catalyzes the reversible oxidation of malate to
CC oxaloacetate (By similarity).
CC -|- CATALYTIC ACTIVITY: (S)-malate + NAD(+) = oxaloacetate + NADH.
CC -|- SIMILARITY: Belongs to the LDH/MDH superfamily. MDH type 3 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC ENBL; BX997699; CAP28418.1; -; Genomic_DNA.
CC HAMAP; MF_00487; -; 1.
CC InterPro; IPR001557; L_LDH_MDH.
CC InterPro; IPR001236; ldh.
CC InterPro; IPR011275; MalatedH_bact.
CC PANTHER; PTHR11540; ldh; 1.
CC Pfam; PF02866; Ldh_1_C; 1.
CC Pfam; PF00056; Ldh_1_N; 1.
CC PIRSF; PIRSF000102; Lac_mal_DH; 1.
CC PRINTS; PRO0086; LLDHGRNASE.
CC TIGRFAMs; TIGR01763; MalatedH_bact; 1.
CC Complete proteome; NAD; Oxidoreductase; Tricarboxylic acid cycle.
FT ACT_SITE 149 149 Proton relay (By similarity).
FT BINDING 83 83 Substrate carboxyl (By similarity).
FT BINDING 89 89 Substrate carboxyl (By similarity).
FT BINDING 152 152 Substrate carboxyl (By similarity).
SQ SEQUENCE 320 AA; 33664 MW; 24C2AF4A2A2C720 CRC64;

Query Match 52.0%; Score 877.5; DB 1; Length 320;
Best Local Similarity 52.8%; Pred. No. 2.8e-58;
Matches 168; Conservative 66; Mismatches 79; Indels 5; Gaps 1;

QY 9 RPKIAMVGSVMGTTAFCLSLRELGDVFLFDVVPNPMGKAMDISHNSSVVDGTITVG 68
DB 3 RKKIALIGSMIGTGLAHITGLKELGDVFLFDIAEGIPQKALDIAESSPVDGFDVSLTG 62
QY 69 SNSYECLKGADVVIITAGITKIPGKSDKESRMDLLFPVNIKIMREYGAATKSYCPNAFVI 128
DB 63 ANSYDVIEGADVIVTAGVARKPG-----NSRDLLGLNLKWMQVQAGIKKYASSAFVI 117
QY 129 NITNPLDVMVAALQESSGLPHRRCGMAGMLDSSRFRMTADKLEVSFRDQGVIVGHG 188
DB 118 CITNPLDAMVWALQKFSGLPTQKVVGMAGILDSSARFRHFLSEEFKISVKDVTAPVLGGHG 177
QY 189 DHMVPLSRYATVNGIPLSEFVKKGWIKQIEVEDDITVQKTKVAGGIVRLGGAGYAPGA 248
DB 178 DSMVPLVRYSTVGISLPDLVGMGTTQEKIDQIIQTRDGGABIVSLTKGSFYAPAA 237
QY 249 SAIQMAESYLKDRKRVVMVCSYQGLQGVQNHLYGVPCVIGGRGVEKIELELTAQRQE 308
DB 238 SAVSMAEAYLKDRKRVVVAAYLSQGVKDYVGVVPWVGAGGVERVIBIDLKBEKAA 297
```

```
QY 309 LOGSIDEVKEMOKATAAL 326
DB 298 FEKSVSQVQLCEACIAV 315

RESULT 11
MDH_BRUME STANDARD; PRT; 320 AA.
AC Q8YJ57;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Malate dehydrogenase (EC 1.1.1.37)
GN Name=mdh; OrderedLocusNames=BMEI0137;
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756888; DOI=10.1073/pnas.221575398;
RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mujer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyrides N.C., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
CC -|- FUNCTION: Catalyzes the reversible oxidation of malate to
CC oxaloacetate (By similarity).
CC -|- CATALYTIC ACTIVITY: (S)-malate + NAD(+) = oxaloacetate + NADH.
CC -|- SIMILARITY: Belongs to the LDH/MDH superfamily. MDH type 3 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC ENBL; AB009456; AAL51319.1; ALT_INIT; Genomic_DNA.
CC PIR; AD3269; AD3269.
CC HSSP; P80040; 1GUY.
CC HAMAP; MF_00487; -; 1.
CC InterPro; IPR001557; L_LDH_MDH.
CC InterPro; IPR001236; ldh.
CC InterPro; IPR011275; MalatedH_bact.
CC InterPro; IPR001252; Mdh_AS.
CC PANTHER; PTHR11540; ldh; 1.
CC Pfam; PF02866; Ldh_1_C; 1.
CC Pfam; PF00056; Ldh_1_N; 1.
CC PIRSF; PIRSF000102; Lac_mal_DH; 1.
CC PRINTS; PRO0086; LLDHGRNASE.
CC TIGRFAMs; TIGR01763; MalatedH_bact; 1.
CC Complete proteome; NAD; Oxidoreductase; Tricarboxylic acid cycle.
FT ACT_SITE 149 149 Proton relay (By similarity).
FT BINDING 83 83 Substrate carboxyl (By similarity).
FT BINDING 89 89 Substrate carboxyl (By similarity).
FT BINDING 152 152 Substrate carboxyl (By similarity).
SQ SEQUENCE 320 AA; 33666 MW; FD73DD8A3654AFB2 CRC64;

Query Match 51.9%; Score 876.5; DB 1; Length 320;
Best Local Similarity 54.2%; Pred. No. 3.4e-58;
Matches 174; Conservative 53; Mismatches 89; Indels 5; Gaps 1;

QY 9 RPKIAMVGSVMGTTAFCLSLRELGDVFLFDVVPNPMGKAMDISHNSSVVDGTITVG 68
DB 3 RKKIALIGSMIGTGLAHITGLKELGDVFLFDIAEGIPQKALDIAESSPVDGFDVSLTG 62
QY 69 SNSYECLKGADVVIITAGITKIPGKSDKESRMDLLFPVNIKIMREYGAATKSYCPNAFVI 128
```

```

Db 63 ANDYAAIEGADVIVTAGVPRKPG-----MSRDDLLGLNLKVMQVGGAGIKKYAPEAFVI 117
Qy 129 NITNPLDVMVAALQESSGLPHRRICGMAGMLDSSRRFRMIADKLEVSFPRDVGQVGVGHG 189
Db 118 CITNPLDAMWALQKFSGLPAHKVGVGMAGVLDLSARFRYFLSEEFNVSVEDVTAFVLGGHG 177
Qy 189 DHMVPLSRVATVNGIPLSEFVKKWKIKOBEVDVIVOKTKVAGGIVRLGQGSAYYAPGA 248
Db 178 DSMVPLARYSTVAGIPLSLDLVKMGWTSQDKLDKIIOQRDGGAEIVGLLTKGSAFYAPAA 237
Qy 249 SAIQMAESYLKDKRVRVMVSCYLOGQGVQVNHVGLVPCVIGGRGVKEIIIELELTAOBRQE 308
Db 238 SAIQMAESYLKDKRVLPAVALQSGQGVKMDYGVVPTVIGANGVERIIEIDLKDEKAQ 297
Qy 309 LQGSIDEVKEKQIAALDAS 329
Db 298 FDKSVASVAGLCEACIGIAPS 318

RESULT 12
MDH_AGRTS
ID MDH_AGRTS STANDARD; PRT; 320 AA.
AC 07CWT; 08UC59;
AT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Malate dehydrogenase (EC 1.1.1.37).
GN Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Agrobacterium.
-OR NCBI_TaxID=176299;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=21608551; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
OKura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
Chapman P., Clendinning J., Deatherage G., Gillet W., Grant C.,
Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
Raymond C., Rouse G., Saenphummachak C., Wu Z., Romero P., Gordon D.,
Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
Nester E.W.;
RA "The genome of the natural genetic engineer Agrobacterium tumefaciens
C58."
RL Science 294:2317-2323 (2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
Ourolo B., Goldman B.S., Cao Y., Akenazi M., Halling C., Mullin L.,
Houmel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
Cielo C., Slater S.;
RA "Genome sequence of the plant pathogen and biotechnology agent
Agrobacterium tumefaciens C58."
RL Science 294:2323-2328 (2001).
CC -1- FUNCTION: Catalyzes the reversible oxidation of malate to
oxaloacetate (By similarity).
CC -1- CATALYTIC ACTIVITY: (S)-malate + NAD(+) = oxaloacetate + NADH.
CC -1- SIMILARITY: Belongs to the LDH/MDH superfamily. MDH type 3 family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AE009211; AAU43620.1; -, Genomic DNA.
CC EMBL; AE008177; AAK88360.1; ALT_INIT; Genomic_DNA.

```

```

DR PIR; AF2900; AF2900.
DR PIR; G97675; G97675.
DR HSP; P80040; IGUY.
DR HAMAP; MF 00487; -.
DR InterPro; IPR001557; L_LDH_MDH.
DR InterPro; IPR001236; Ldh_-.
DR InterPro; IPR011275; MalateDH_bact.
DR InterPro; IPR001252; Mdh_AS.
DR PANTHER; PTHR11540; Ldh; 1.
DR Pfam; PF02866; Ldh_1_C; 1.
DR Pfam; PF00056; Ldh_1_N; 1.
DR PIRSF; PIRSF000102; Lac_mal_DH; 1.
DR PRINTS; PRO0086; LLDHDEGNASE.
DR TIGR; TIGR01763; MalateDH_bact; 1.
KW Complete proteome; NAD; Oxidoreductase; Tricarboxylic acid cycle.
FT ACT_SITE 149 176 proton relay (By similarity).
FT ACT_SITE 176 176 proton relay (By similarity).
FT BINDING 83 83 Substrate carboxyl (By similarity).
FT BINDING 89 89 Substrate carboxyl (By similarity).
FT BINDING 152 152 Substrate carboxyl (By similarity).
SQ SEQUENCE 320 AA; 33537 MW; 3073A7C3BC0FCE5E CRC64;

Query Match 51.7%; Score 873.5; DB 1; Length 320;
Best Local Similarity 53.2%; Pred. No. 5.7e-58;
Matches 167; Conservative 60; Mismatches 82; Indels 5; Gaps 1;

Qy 9 RPKIAMVSGMIGTMAFLCSLRELGDVVLFDVVPNPMGKAMDISHNSVVDITGVY 68
Db 3 RKKIALIGSGMIGTTLAHLASLAKELGDIVLFDIADGIPQKGLDIAQSGPVEGFNAKLSG 62
Qy 69 SNSYECLKGDVVIITAGITKIPKSDKEWSRDLPLVNIKIMREVCAAIKSYCPNAFVI 128
Db 63 ASDYAAIEGADVIVTAGVPRKPG-----MSRDDLLGLNLKVMQVGGAGIKKYAPEAFVI 117
Qy 129 NITNPLDVMVAALQESSGLPHRRICGMAGMLDSSRRFRMIADKLEVSFPRDVGQVGVGHG 189
Db 118 CITNPLDAMWALQKFSGLPKNVGMAGVLDLSARFRYFLSEEFNVSVEDVTAFVLGGHG 177
Qy 189 DHMVPLSRVATVNGIPLSEFVKKWKIKOBEVDVIVOKTKVAGGIVRLGQGSAYYAPGA 248
Db 178 DTWVPLARYSTVAGVPLTDLVKMGWLTAELEQIQTRDGGAEIVGLLTKGSAFYAPAA 237
Qy 249 SAIQMAESYLKDKRVRVMVSCYLOGQGVQVNHVGLVPCVIGGRGVKEIIIELELTAOBRQE 308
Db 238 SAIQMAESYLKDKRVLPAVALQSGQGVKMDYGVVPTVIGANGVERIIEIDLKDEKAQ 297
Qy 309 LQGSIDEVKEKQIA 322
Db 298 FDKSVASVAGLCEA 311

RESULT 13
MDH_RHIME
ID MDH_RHIME STANDARD; PRT; 320 AA.
AC Q9EYU6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Malate dehydrogenase (EC 1.1.1.37).
GN Name=mdh; OrderedLocustNames=R03056; ORFNames=SMC02479;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Dymov S.I., Meek D.J., Driscoll B.T.;
RT "Isolation of a malate dehydrogenase mutant and genes encoding a
RT putative TCA cycle operon of Sinorhizobium meliloti."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=1021;

```

[illegible]

RESULT 14
MDH BARQU

ID	MDL_BARQU	STANDARD;	PRT;	320 AA.
AC	Q6FYD0;			
DT	25-OCT-2004	(Rel. 45, Created)		
DD	25-OCT-2004	(Rel. 45, Last sequence update)		
DT	13-SEP-2005	(Rel. 48, Last annotation update)		
DE	Malate dehydrogenase [EC 1.1.1.37].			
GN	Name=mdh; OrderedLocusNames=BQ13450;			
OS	Bartonella quintana (Rochalimaea quintana).			
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;			
OC	Bartonellaceae; Bartonella.			
OX	NCBI_TaxID=903;			
RN	[1]			
RC	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].			
RC	STRAIN=Toulouse;			
RX	PubMed=15210978; DOI=10.1073/pnas.0305659101;			
RA	Aismark U.C.M., Frank A.C., Karlberg E.O., Legault B.-A., Ardelet D.H.,			
RA	Canbaeck B., Eriksson A.-S., Naeslund A.K., Handley S.A., Huvet M.,			
RA	La Scola B., Holmberg M., Andersson S.G.E.;			
RT	"The louse-borne human pathogen Bartonella quintana is a genomic derivative of the zoonotic agent Bartonella henselae.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 101:9716-9721(2004).			
CC	-!- FUNCTION: Catalyzes the reversible oxidation of malate to oxaloacetate (By similarity).			
CC	-!- CATALYTIC ACTIVITY: (S)-malate + NAD(+)=oxaloacetate + NADH.			
CC	-!- SIMILARITY: Belongs to the LDH/MDH superfamily. MDH type 3 family.			
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.			
DR	EMBL; BX997700; CAE26803.1; -; Genomic_DNA.			
DR	HMAP; MF_00487; -; 1			
DR	InterPro; IPR001088; Glyco_hydro_4.			
DR	InterPro; IPR001557; L_LDH_MDH.			
DR	InterPro; IPR001236; ldh.			
DR	InterPro; IPR011275; MalatedH_bact.			
DR	InterPro; IPR001252; Mdh_AS.			
DR	PANTHER; PTHR11540; ldh; 1.			
DR	Pfam; PF02866; Ldh_1_C; 1.			
DR	Pfam; PF00056; Ldh_1_N; 1.			
DR	PRINTS; PRSF000102; Lac_mal_DH; 1.			
DR	PRINTS; PRO0086; LLHDHGNASE.			
DR	ProDom; PD006892; Glyco_hydro_4; 1.			
DR	TIGRFAMs; TIGR01763; MalatedH_bact; 1.			
KW	Complete proteome; NAD; Oxidoreductase; Tricarboxylic acid cycle.			
FT	ACT SITE 149 176 176 Proton relay (By similarity).			
FT	ACT SITE 176 176 Proton relay (By similarity).			
FT	BINDING 83 83 Substrate carboxyl (By similarity).			
FT	BINDING 89 89 Substrate carboxyl (By similarity).			
FT	BINDING 152 152 Substrate carboxyl (By similarity).			
SQ	SEQUENCE 320 AA; 33935 MW; 5ABDB8A64348D544 CRC64;			
	Query Match			
	Best Local Similarity 51.6%; Score 870.5; DB 1; Length 320;			
	Matches 170; Conservative 62; Mismatches 81; Indels 5; Gaps 1;			
Qy	9 RPKIAMVSGMIGTWAFLCSBELGDVLFDVVVPNMPGMKAMDISHNSVVDGTGTVYG 68 : : : : :			
Dd	3 RKKIATGSGMIGTTAHMTIKGLKELGDVLFDIEEGPQGKALDIASFSFVDGDFVSLTG 62 : : : : :			
Qy	69 SNSVECLKGADVITTAGITKIPKSDKEWSRDLLPVNTIKIMREVGAAIKSYCPNAFVI 128 : : : : :			
Dd	63 ANYTEATEGADVIVITAGVARKEP-----MSRDDLLGINLKVMQGVAGIKKYASSAFVI 117 : : : : :			
Qy	129 NITNPLDMVAALQESSGLPHHRICGMAGMLDSRRFRMIADKLVSPPRDQMGVIGHG 188 : : : : :			
Dd	118 CIINPLDMVMVALQKESGLPKQKVGVAGVLDARSFRYFLSKFKSVKDVDAFVLGGHG 177 : : : : :			
Qy	189 DHMVPLSRYSATVNGIPISEFVKGWTKQEVEDDIVOKTKVAGEIIVRLCQSAYYPAGA 248 : : : : :			
Dd	178 DSMVPLVRYSVTGGISLPDIIVKNGWTQTQEKIDIOIIQRVNRNGAEIIVGLLKTGSFAFPAA 237 : : : : :			

Qy	249	SAIQWAEASYLKDKRKVMWCVCYLQCYQYGVNHYLGVPCVIGRGVKEKIELELTAQERQE	308
Db	238	SAISMAEAYLKDKRKVPVATYLSGEYGVKDTYGVVVLGAGGVVERVIEIDLKKERSA	297
Qy	309	LOGSIDVEKEMOKATAAL	326
Db	298	FQSVNAVKKUCEACIAI	315
RESULT 15			
MDH_BRUAB			
ID	MDH_BRUAB	STANDARD;	PRT; 320 AA.
AC	Q57AXI;		
DT	13-SEP-2005 (Rel. 48, Created)		
DT	13-SEP-2005 (Rel. 48, Last sequence update)		
DT	13-SEP-2005 (Rel. 48, Last annotation update)		
DE	Malate dehydrogenase (EC 1.1.1.37)		
GN	Name=mdh; OrderedLocusNames=BruAb1_1903;		
OS	Brucella abortus.		
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;		
OC	Brucellaceae; Brucella.		
OX	NCBI_TaxID=235;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].		
RC	Strain=9-941 / Biovar 1;		
RX	PubMed=15805518; DOI=10.1128/JB.187.8.2715-2726.2005;		
RA	Halling S.M., Peterson-Burch B.D., Bricker B.J., Zuerner R.L.,		
RA	Qing Z., Li L., Kapur V., Alt D.P., Olsen S.C.;		
RT	"Completion of the genome sequence of Brucella abortus and comparison		
RT	to the highly similar genomes of Brucella melitensis and Brucella		
RT	suis.";		
RL	J. Bacteriol. 187:2715-2726(2005).		
CC	-1- FUNCTION: Catalyzes the reversible oxidation of malate to		
CC	oxaloacetate (By similarity).		
CC	-1- CATALYTIC ACTIVITY: (S)-malate + NAD(+) = oxaloacetate + NADH.		
CC	-1- SIMILARITY: Belongs to the LDH/MDH superfamily. MDH type 3 family.		
CC			
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use as long as its content is in no way modified and this statement is not		
CC	removed.		
CC			
CC	EMBL; AE017223; AAX75213.1; -; Genomic_DNA.		
DR	HMAP; MF 00487; -; 1.		
DR	InterPro; IPR001557; L_LDH_MDH.		
DR	InterPro; IPR001236; ldn.		
DR	InterPro; IPR011275; MalatedH_bact.		
DR	PANTHER; PTHR11540; ldn; 1.		
DR	Pfam; PF02866; Ldh_1_C; 1.		
DR	Pfam; PF00056; Ldh_1_N; 1.		
DR	PfRSF; PFRSF00102; Lac mal DH; 1.		
DR	PRINTS; PR00086; LLDHDRGNASE.		
DR	TIGRFAMs; TIGR01763; MalatedH_bact; 1.		
KW	Complete proteome; NAD; Oxidoreductase; Tricarboxylic acid cycle.		
FT	ACT_SITE 149 149		
FT	ACT_SITE 176 176		
FT	BINDING 83 83		
FT	BINDING 89 89		
FT	BINDING 152 152		
FT	SEQUENCE 320 AA; 33704 MW; C212BA88F0241677 CRC64;		
Qy	9	RPKIAMVSGMIGGTMAFLCRLRELGDVVLFDVVPNMPGKAMDISHNSVVDITGITVYG	68
Db	3	RNKALIGSMIGGTLAHLAGLKGELGDVVLFDIAEGTPQGLDIAESSPVDGPAKFTG	62
Qy	69	SNSYECLKGADVITITAGITKIPGKSDEKWSRMDLLPVNIKIMREVGAAIKSYCPNAPVI	128
Query Match			
Best Local Similarity 51.5%; Score 868.5; DB 1; Length 320;			
Matches 173; Conservative 53; Mismatches 90; Indels 5; Gaps 1			

THIS PAGE BLANK (uspto)

Result No.	Query #			DB	ID	Description
	Score	Match	Length			
1	876.5	51.9	326	2	AD3269	malate dehydrogena
2	873.5	51.7	320	2	AF2300	malate dehydrogena
3	873.5	51.7	352	2	G97675	malate dehydrogena
4	850.5	50.4	320	2	E87702	malate dehydrogena
5	837.5	49.6	314	2	H97764	malate dehydrogena
6	826.5	49.0	314	2	A71695	malate dehydrogena
7	740	43.8	312	2	I40383	malate dehydrogena
8	729	43.2	312	2	S61213	malate dehydrogena
9	709	42.0	314	2	F84044	malate dehydrogena
10	634	37.6	324	2	S75735	probable malate de
11	605.5	35.9	335	2	E70453	malate dehydrogena
12	569	33.7	334	2	D70444	malate dehydrogena
13	566.5	33.6	284	2	AC2346	malate dehydrogena
14	560.5	33.2	313	2	E72555	probable malate de
15	478	28.3	314	2	H86671	L-lactate dehydrog
16	465.5	27.6	313	2	AC1463	L-lactate dehydrog
17	463.5	27.5	313	2	AC1101	L-lactate dehydrog
18	460	27.3	321	2	E96949	L-lactate dehydrog
19	455	27.0	318	1	DEBSLM	L-lactate dehydrog
20	454	26.9	313	2	E96932	L-lactate dehydrog
21	440.5	26.1	353	2	F71441	L-lactate dehydrog
22	440	26.1	318	2	S08182	L-lactate dehydrog
23	439.5	26.0	319	2	S36863	L-lactate dehydrog
24	434.5	25.7	312	2	H64250	L-lactate dehydrog
25	431.5	25.6	326	1	DEUBLA	L-lactate dehydrog
26	431	25.5	319	2	S08183	L-lactate dehydrog
27	428	25.4	317	2	B29704	L-lactate dehydrog
28	424.5	25.1	317	2	S00019	L-lactate dehydrog
29	423.5	25.1	320	2	JQ0181	L-lactate dehydrog

```
RESULT 2
AF2900
malate dehydrogenase mdh [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AF2900
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.;
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Kreepan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AF2900
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-320 <R>
A;Cross-references: UNIPROT:Q8UC59; UNIPARC:UPI000016474B; GB:AE008688; PIDN:AA443620.1;
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: mdh
A;Map position: circular chromosome
C;Superfamily: L-lactate dehydrogenase

Query Match 51.7%; Score 873.5; DB 2; Length 320;
Best Local Similarity 53.2%; Pred. No. 1.3e-59;
Matches 167; Conservative 60; Mismatches 82; Indels 5; Gaps 1;

Qy 9 RPKIAMVSGMIGTMAFLCSRLGDLVFLFDVVPNPMGKAMDISHNSSVVDGTITVYG 68
Db 3 RKXIALIGSGMIGTTLAHLASLKLGDIVLFDIADGIPQKGLDIAQSGPVEGFNAKLSG 62

Qy 69 SNSYECLKGADVVIITAGITKIPGSKDKWSRMDLLPVNIKIMREVGAAIKSYCPNAPVI 128
Db 63 ASDYAAIEGADVCIIVTAGVARKPG-----MSRDDLLGLNLKVMQEQGAGIKKYAPNAPVI 117

Qy 129 NITNPLDMVVAALQESSGLPHHRICGMAGMLDSSRRFRMIADKLEYSRDPVQGMVIGVHG 198
Db 118 CITNPLDMVVAALQESSGLPHHRICGMAGMLDSSRRFRMIADKLEYSRDPVQGMVIGVHG 177

Qy 189 DHMVPLSRVATVNGIPLSEFVKKGWIKQBEVDDIVQKTKVAGGEIVRLLLGQGSAYYAPGA 248
Db 178 DMVPMVRHSTVGGGLPLPLVKQWLSQDKLDAIVERTRKGGGIEIVALLKTSYAFAPAE 237

Qy 249 SAIQMAESYLKDKRVMVCSYLQGGYQVQNHLYGVPVIGRGVKEKIELELTAQERQE 308
Db 298 FOKSVGAVAGLCEA 311

RESULT 3
G97675
malate dehydrogenase [imported] - Agrobacterium tumefaciens (strain C58, Cerson)
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: G97675
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: G97675
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-352 <R>
A;Cross-references: UNIPROT:Q8UC59; UNIPARC:UPI000000D1F8D; GB:AE007869; PIDN:AAK88360.1;
C;Genetics:
A;Gene: AGR_C_4782
A;Map position: circular chromosome
```

C;Superfamily: L-lactate dehydrogenase

```
Query Match 51.7%; Score 873.5; DB 2; Length 352;
Best Local Similarity 53.2%; Pred. No. 1.5e-59;
Matches 167; Conservative 60; Mismatches 82; Indels 5; Gaps 1;

Qy 9 RPKIAMVSGMIGTMAFLCSRLGDLVFLFDVVPNPMGKAMDISHNSSVVDGTITVYG 68
Db 35 RKXIALIGSGMIGTTLAHLASLKLGDIVLFDIADGIPQKGLDIAQSGPVEGFNAKLSG 94

Qy 69 SNSYECLKGADVVIITAGITKIPGSKDKWSRMDLLPVNIKIMREVGAAIKSYCPNAPVI 128
Db 95 ASDYAAIEGADVCIIVTAGVARKPG-----MSRDDLLGLNLKVMQEQGAGIKKYAPNAPVI 149

Qy 129 NITNPLDMVVAALQESSGLPHHRICGMAGMLDSSRRFRMIADKLEYSRDPVQGMVIGVHG 198
Db 150 CITNPLDMVVAALQESSGLPHHRICGMAGMLDSSRRFRMIADKLEYSRDPVQGMVIGVHG 209

Qy 189 DHMVPLSRVATVNGIPLSEFVKKGWIKQBEVDDIVQKTKVAGGEIVRLLLGQGSAYYAPGA 248
Db 210 DTWVPLARYSTVGGVPLTDLVKGWLTAELEQIIQTRDGGAEIVLLKTSYAYAPAA 269

Qy 249 SAIQMAESYLKDKRVMVCSYLQGGYQVQNHLYGVPVIGRGVKEKIELELTAQERQE 308
Db 270 SAIQMAESYLKDKRVLPAHAHLSGQYGVDDMVGVPTIIGAGGIERVIEIENKEEAA 329

Qy 309 LQGSIDEVKEMQKA 322
Db 330 FOKSVGAVAGLCEA 343
```

RESULT 4

E87702

malate dehydrogenase [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C;Accession: E87702

R;Niernan, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eien, J.; Heidelberg, J.

n, J.; Laub, M.R.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: E87702

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-320 <STO>

A;Cross-references: UNIPROT:Q9A2B1; UNIPARC:UPI00000C7B5D; GB:AE005673; NID:g13425411; PJ

C;Genetics:

A;Gene: CC3655

C;Superfamily: L-lactate dehydrogenase

```
Query Match 50.4%; Score 850.5; DB 2; Length 320;
Best Local Similarity 51.1%; Pred. No. 7.7e-58;
Matches 164; Conservative 60; Mismatches 92; Indels 5; Gaps 1;
```

Qy 9 RPKIAMVSGMIGTMAFLCSRLGDLVFLFDVVPNPMGKAMDISHNSSVVDGTITVYG 68

Db 3 RAKIALIGAMIGGTTLAHLAAREELGDVILFDIAEGTPQKGLDIAEASAVFGKDVALKG 62

Qy 69 SNSYECLKGADVVIITAGITKIPGSKDKWSRMDLLPVNIKIMREVGAAIKSYCPNAPVI 128

Db 63 ANDYADIGADVCIIVTAGVPRKPG-----MSRDDLLGLNLKVMKAVGEGIKAHAPNAPVI 117

Qy 129 NITNPLDMVVAALQESSGLPHHRICGMAGMLDSSRRFRMIADKLEYSRDPVQGMVIGVHG 198

Db 118 CITNPLDMVVAALQESSGLPHHRICGMAGMLDSSRRFRMIADKLEYSRDPVQGMVIGVHG 177

Qy 189 DHMVPLSRVATVNGIPLSEFVKKGWIKQBEVDDIVQKTKVAGGEIVRLLLGQGSAYYAPGA 248

Db 178 DMVPMVRHSTVGGGLPLPLVKQWLSQDKLDAIVERTRKGGGIEIVALLKTSYAFAPAE 237

Qy 249 SAIQMAESYLKDKRVMVCSYLQGGYQVQNHLYGVPVIGRGVKEKIELELTAQERQE 308

Db 238 SAJAMATSYLKDKRVRPCATYLTGQGLNDLYGVVPVWVGAGAEKIVFETNDEKAM 297
QY 309 LOGSIDEVEMQKAIKALDAS 329
Db 298 FAKSVESVKLMEACKAIDSS 318
RESULT 5
H97764
malate dehydrogenase (EC 1.1.1.37) [imported] - Rickettsia conorii (strain Malish 7)
C:Species: Rickettsia conorii
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: H97764
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R
Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:2142074; PMID:11557893
A:Accession: H97764
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-314 <KUR>
A:Cross-references: UNIPROT:Q921A0; UNIPARC:UPI000012EE50; GB:A5006914; PIDN:AAL03058.1.
C:Genetics:
C:Superfamily: L-lactate dehydrogenase
C:Keywords: oxidoreductase
Query Match 49.68; Score 837.5; DB 2; Length 314;
Best Local Similarity 49.8%; Pred. No. 7.5e-57;
Matches 156; Conservative 72; Mismatches 80; Indels 5; Gaps 1;
QY 11 KIAMVSGMIGGTMAFLCSRLRELDVDFVVPNMPMGKAMDISHNSVVDGTGTVVGSN 70
Db 6 KISLISGNGTGLLAHLISLRELDVDFVTEGVPGKALDLMQAGTIGSDIKIKGTN 65
QY 71 SYECLGADVVIITAGITKTPGSKDSEWMDLLPVNIKIMREVGAIAKSCYCNAPVNI 130
Db 66 DYKDIIEGSDAIIITAGLPKPG-----MSREDLISINTGIMKTVAANVKYAPDAFVIV 120
QY 131 TNPDLVMAALQESSGLPHRHICGMAGMLDSSRRFRMIADKLEVPDVGWGVHGDH 190
Db 121 TNPDLVMAVYMLKESGLPHNKGVLDSRRFLAEFKVSVNVMNVLGGHGA 180
QY 191 MVLPSRYATVNGIPLSFVFKGMKQBEVDIVOKTKVAGGEIVRLGOGSAYYAPGASA 250
Db 181 MVPLARYSTISGVPIPLIKWGLSSNENIEKIDRTNGGGEIVALLKTGSAYYAPASA 240
QY 251 IQMAESYLKDKRKMVWCSCYLGQYGVNHYLGVPCVIGRGVEKIIIELELTAEQELQ 310
Db 241 IEMLESYLDKQRIILTCAAHLQGEYGVHDLVGVPIMGKGVLRVIELQTAEEKALFD 300
QY 311 GSIDEVEMQKAI 323
Db 301 KSVGEGVKLIETI 313
RESULT 6
A71695
malate dehydrogenase (mdh) RP376 - Rickettsia prowazekii
C:Species: Rickettsia prowazekii
C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004
C:Accession: A71695
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark,
Nature 396, 133-140, 1998
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A:Reference number: A71630; MUID:99039499; PMID:9823893
A:Accession: A71695
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-314 <AND>
A:Cross-references: UNIPROT:Q92DF3; UNIPARC:UPI000012EE51; GB:AJ235271; GB:AJ235269; NID:
A:Experimental source: strain Madrid E

C:Genetics:
A:Gene: mdh; RP376
C:Superfamily: L-lactate dehydrogenase
Query Match 49.0%; Score 826.5; DB 2; Length 314;
Best Local Similarity 49.8%; Pred. No. 5.3e-56;
Matches 157; Conservative 70; Mismatches 81; Indels 7; Gaps 2;
QY 5 EKNTRPKIAMVSGMIGGTMAFLCSRLRELDVDFVVPNMPMGKAMDISHNSVVDGTGI 64
Db 2 KKN--EKISLISGNGTGLLAHLISLRELDVDFVTEGVPGKALDLMQAAATIEGSDI 59
QY 65 TVYGSNSYELCKGADVVIITAGITKTPGSKDSEWMDLLPVNIKIMREVGAIAKSCYCN 124
Db 60 KIKGTNDYRDIEGSDAIIITAGLPKPG-----MSRDDLSISVNTKMKDVAQNIKYAQN 114
QY 125 AFVINITNPLDVMAALQESSGLPHRHICGMAGMLDSSRRFRMIADKLEVPDVGWGMV 184
Db 115 AFVIVITNPLDVMAVYMLKESGLPHNKGVLDSRRFLAEFKVSVNVMNVLGGHGA 174
QY 185 GVHGDHMPVLSRYATVNGIPLSFVFKGMKQBEVDIVOKTKVAGGEIVRLGOGSAYY 244
Db 175 GGHGDTMVLPLRYSTISGVPIPLIKWGLSSNENIEKIDRTNGGGEIVALLKTGSAYY 234
QY 245 APCASAIQMAESYLKDKRKMVWCSCYLGQYGVNHYLGVPCVIGRGVEKIIIELELTAE 304
Db 235 APAASAIAMLESYLDKQRIILTCAAHLQGEYGVHDLVGVPIMGKGVLRVIELQTEE 294
QY 305 ERQELQGSIDEVEMK 319
Db 295 EKILFYKSVTEVKKL 309
RESULT 7
I40383
malate dehydrogenase (EC 1.1.1.37) citH - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004
C:Accession: I40383; B69600
R:Jin, S.; Sonenshein, A.L.
J. Bacteriol. 176, 4669-4679, 1994
A:Title: Identification of two distinct Bacillus subtilis citrate synthase genes.
A:Reference number: I40379; MUID:94321340; PMID:8045898
A:Accession: I40383
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-312 <RES>
A:Cross-references: UNIPROT:P49814; UNIPARC:UPI00000608B9; EMBL:U05257; NID:g1045295; PII
R:Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertec
C: Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabbret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
tech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: B69600
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-312 <KUN>
A:Cross-references: UNIPARC:UPI00000608B9; GB:Z99118; GB:AL009136; NID:g2635200; PIDN:CA
A:Experimental source: strain 168
C:Genetics:
C:Superfamily: L-lactate dehydrogenase

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena* sp. PCC 7120
A;Accession number: AB1807; MUID:21595285; PMID:11759840
A;Reference: AC2346
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-284 <KUR>
A;Cross-references: UNIPROT:Q8YP78; UNIPARC:UPI00000CEABC; GB:BA000019; PIDN:BAB76021.1;
C;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr4322
C;Superfamily: L-lactate dehydrogenase

Query Match	33.6%	Score 566.5;	DB 2;	Length 284;
Best Local Similarity	42.0%	Pred. NO. 4.4e-36;		
Matches 124; Conservative	54;	Mismatches 90;	Indels 27;	Gaps 5;

Qy	36	VULFDVVPNPMGKAMDI-----SHNSVVDTGITYGVSNSYECILGKADVILITAGIT	88
Db	1	MVLVDIVEGHPOGLADLLLEARGIEUHNQII-----GTNNYADTSGSIVVITAGFP	53
Qy	89	KIPGSKDEKWSRMDLLPVMIKMKREVGAAIKSYCPNAPFINITNPLDVMVAALQESSGLP	148
Db	54	RKPG-----MSRDDLLRTNAKIVIEAAKQAIAYSPTAIFVVTNPLDVMYLAWEANGLP	108
Qy	149	HHRICGAGMLDSSRFRPMIADKLEVSPRDVQGVIGVGHGDHWPPLSYATVNGIPIUSEF	208
Db	109	RNRIMGAGVLDARSPTETIALEGLVPADVKAVMYLGSHGLMWPLSSYATVNGIPIQTQ	168
Qy	209	VKKGWIKQEBVDDIVOKTKVAGEIIVLLLGGSAYYAPCASAIQMAESVLDKDRKVMVCS	268
Db	169	-----LDVNTIELVSRTRNGAEIVELMQTGAFAPASATSLMVESILLNQRLPVS	228
Qy	269	CYLQGYGVQNHVLPVPCVIGRGVEKIIIELTATQERQELQGSIDVEKEMQKAI	323
Db	224	IYLOGEYLDKDVVIGPCRLGLNGTSEVLENLSDSREALHIS---AKSVOKNI	275

RESULT 14
E72655

Probable malate dehydrogenase APE0672 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: E72655
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahashi, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kawarabayashi, Y. 1999
A:Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix strain K1
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: E72655
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-313 <KAW>
A:Cross-references: UNIPROT:Q9VEA1; UNIPARC:UPI0000005DC3A; DDBJ:AP000060; NID:G5104189;
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0672
C:Superfamily: L-lactate dehydrogenase

Query Match 33.2%; Score 560.5; DB 2; Length 313;
Best Local Similarity 35.9%; Pred. No. 1.4e-35;
Matches 113; Conservative 72; Mismatches 119; Indels 11; Gaps 3;

Qy 9 RPKIAMVGSWIGCTMAFLCSLRELGDVLFVDVVPNPMGKAMDISHNSSVVDTGITYG 68

DD 4 VFTIIILGAGKGVMAIAVFLMWMKGIIDULLLIJAKIPKPFQSGAADDUTAHAAWAEJGVJIKISG 83

QV 69 SNSVECLKGADWIIITAGITKIPKSKDKWESRMDDLTPWIKIMREVGNAIKSYCYPNAFVI 128

Db 64 SNSYEDMRGSDIVLTAGIRKPG-----MTREQLEAANTMADLAEIKIYAKAIV 118

QY	129	N	I	N	P	L	D	V	M	A	A	L	S	S	G	L	P	H	H	R	I	C	M	A	G	M	L	D	S	S	R	F	R	M	I	A	D	K	L	V	S	P	R	D	V	G	M	V	I	G	V	H	G		
DP	119	I	T	T	N	P	D	M	A	T	T	V	M	Y	M	K	T	G	F	P	R	E	V	T	G	F	S	I	L	D	S	A	R	M	A	V	T	S	O	K	I	G	V	S	F	K	S	V	N	A	I	V	L	M	H

Qy	189	DHNVPLSRYATVNGIPISFEFVKFGWIKQEEVDDIVQTKVAGGEIVRLGGQSAIYPAGA	248
Db	179	QKMFPPRLSSVGGVPLEHLSK-----EETEEVSVETNAGAKITELRGY--SSNYGPAA	232
Qy	249	SAIQMASYSYLKDRKVMVCSYLGQGYGVQNHLYLGPVCIIGRGVGEKIIIELTQAQRQE	308
Db	233	GLVLTVEAIRKDSKRIYPSLYLQGEYGYNDIVAEVPAVIGKSGIERIIELPLTEDEKRR	292
Qy	309	LQGSIDEVKEMQAKI	323
Db	293	FDEAVOAVKKLVETL	307

RESULT 15
H86671

L-lactate dehydrogenase (EC 1.1.1.27) [imported] - Lactococcus lactis subsp. lactis (str
 C:Species: Lactococcus lactis subsp. lactis
 C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
 C:Accession: H86671
 R:Botolin, A.; Winkler, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
 Genome Res. 11, 731-753, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp
 A:Reference number: H86625; MUID:21235186; PMID:1137471
 A:Accession: H86671
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-314 <STO>
 A:Cross-references: UNIPROT_Q9CII4; UNIPARC:UPI000012E2D6; GB:AE005176; PID:gl2723247; PJ
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: ldhB
 C:Superfamily: L-lactate dehydrogenase
 C:Keywords: Oxidoreductase

Query Match 28.3%; Score 478; DB 2; Length 314;
Best Local Similarity 32.8%; Pred. No. 3.1e-29;
Matches 105; Conservative 79; Mismatches 126; Indels 10; Gaps 5;

QY 6 KNTRPKIAMVSGMIGGTMAF-LCSLRELGDVVLFDVVPNMPMGKAMDISHNSSVVDGTGI 64

Db 2 KITSRKVVIGTGFVGTSAYSMINQGLVNLVLIDVNQDKAEGEALDLLDGVSWGQENV 61

QY 65 TVYGSNSYECLKGADVWIIITAGITKIPGKSDKEWSRMDLLPVNIKIMREVGAAIKSYCPN 12

Db 62 IVRAGDYKDC-KNADI VVVVTAGVNQKPGQ-----SRLDLVNTNAKIMRSIVTQVMDSGFD 11

QY 125 AFVINITNPDLVNMVAALQSSGLPHHRCICGMAGMLDSSRRFRMIADKLEVSPRDVQGMVI 184

QV 185 GVTHGDHNVPLSRVATVNGIPLSEF-VKKGWIKQEEVDDIVQKTKVAGGEIVRLLGGSAY 243

176 GEHGDSEVAVWSHTTVGCKPILEFIVKXKKIGVEDLSNLSKVKNAAEIID--KKQATY 233
Db

QY 244 YAPGASAIQMAESYLKDRKRVMVCSYLGQYGVQNHVGLVPCVIGRGVEKIIIELELTA 303

Db 234 YGIGMSTARIVKAILNNEQAILPVSAYLGRGEYQEGVFTGVPSIVNQNGVREIIELNIDA 293

QY 304 QERQELQSGSIDEVKEMQKAI 323

db 294 YEKKQFEKSVSQLKEVIESI 313

Search completed: March 2, 2006, 19:40:42
Job time : 25 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 19:37:51 ; Search time 185 Seconds
(without alignments)
783.757 Million cell updates/sec

Title: US-09-390-846-2
Perfect score: 1688
Sequence: 1 MAVFEKTRPKIAVGSGLI.....GSIDEVKEMQKAAALDASK 330

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1685	99.8	330	2 AAW11476	Aaw11476 Eimeria 1
2	880.5	52.2	320	8 ADS28805	Ads28805 Bacterial
3	873.5	51.7	319	8 ADS22408	Ads22408 Bacterial
4	873.5	51.7	320	8 ADS25456	Ads25456 Bacterial
5	873.5	51.7	320	8 ADS25841	Ads25841 Bacterial
6	862.5	51.1	311	8 ADS42318	Ads42318 Bacterial
7	862	51.1	322	8 ADS23027	Ads23027 Bacterial
8	850.5	50.4	320	8 ADS28046	Ads28046 Bacterial
9	848.5	50.3	320	8 ADN25360	Adn25360 Bacterial
10	847.5	50.2	320	8 ADS21490	Ads21490 Bacterial
11	828	49.1	315	2 AAR62605	Aar62605 P. falcip
12	828	49.1	315	2 AAY01690	Aay01690 A. Plasmid
13	740	43.8	312	8 ADS44807	Ads44807 Bacterial
14	709	42.0	314	8 ADS28362	Ads28362 Bacterial
15	706	41.8	329	2 AAR94013	Aar94013 Heat resi
16	697	41.3	312	8 ADS27485	Ads27485 Bacterial
17	694	41.1	289	9 ABM94287	Abm94287 M. xanthu
18	689	40.8	285	8 ADS30094	Ads30094 Bacterial
19	672	39.8	309	8 ADN26392	Adn26392 Bacterial
20	669.5	39.7	319	8 ADS24552	Ads24552 Bacterial
21	652.5	38.7	304	8 ADS21368	Ads21368 Bacterial
22	649.5	38.5	325	8 ADS23187	Ads23187 Bacterial
23	641	38.0	325	5 ABP40688	Abp40688 Staphyloc
24	641	38.0	325	8 ADS07962	Ads07962 Staphyloc

25	635	37.6	304	8	ADS22139	Ads22139 Bacterial
26	635	37.6	317	8	ADS30817	Ads30817 Bacterial
27	634	37.6	324	8	ADN20169	Adn20169 Bacterial
28	628	37.2	299	8	ADN27111	Adn27111 Bacterial
29	626.5	37.1	305	8	ADS21934	Ads21934 Bacterial
30	625.5	37.1	309	8	ADS29825	Ads29825 Bacterial
31	605.5	35.9	335	8	ADN17485	Adn17485 Bacterial
32	569	33.7	334	8	ADN17472	Adn17472 Bacterial
33	560.5	33.2	313	8	ADS44426	Ads44426 Bacterial
34	529	31.3	197	8	ADS25293	Ads25293 Bacterial
35	478	28.3	314	5	ABBS3683	Abbs3683 Lactococc
36	478	28.3	314	8	ADS29297	Ads29297 Bacterial
37	464	27.5	324	7	ADC96859	Adc96859 E. faeciu
38	463.5	27.5	313	5	ABBA7671	Abba7671 Listeria
39	462	27.4	312	8	ADS27646	Ads27646 Bacterial
40	460	27.3	321	8	ADOS9775	Ados9775 B. subtil
41	460	27.3	321	8	ADS44643	Ads44643 Bacterial
42	458	27.1	312	8	ADS27670	Ads27670 Bacterial
43	455	27.0	318	6	ABR82282	Abrr82282 B. megate
44	440.5	26.1	353	3	AAG30183	Aag30183 Arabidops
45	439.5	26.0	319	8	ADN20416	Adn20416 Bacterial

ALIGNMENTS

RESULT 1
AAW11476
ID AAW11476 standard; protein; 330 AA.

XX AC AAW11476;
XX AC AAW11476;
DT 17-OCT-2003 (revised)
DT 29-APR-1997 (first entry)
XX Eimeria lactate dehydrogenase.
XX Eimeria lactate dehydrogenase.
XX Lactate dehydrogenase; LDH; coccidiosis; vaccine; vector.
XX Eimeria acervulina; strain Houghton.
OS Eimeria acervulina; strain Houghton.
PN AU9656287-A.
PD 16-JAN-1997.
XX 02-JUL-1996; 96AU-00056287.
XX 03-JUL-1995; 95EP-00201801.
XX (ALKU) AKZO NOBEL NV.
XX Kok JJ, Van Den Boogaart P, Vermeulen AN;
XX WPI; 1997-109375/11.
XX N-PSDB; AAT51370.
XX Eimeria lactate dehydrogenase protein - used for prodn. of vaccines
XX against coccidiosis in poultry.
PS Claim 3; Page 20-22; 30pp; English.
XX The 37 kDa lactate dehydrogenase (LDH) (AAW11476) of Eimeria acervulina
XX schizonts, or immunologically active portions of it, can be used for
XX prodn. of vaccines against coccidiosis in poultry. The LDH can be
XX isolated from schizonts or produced in host cells or organisms
XX transformed with recombinant vectors including LDH nucleic acids (see
XX also AAT51370). Live viral vaccines can also be produced. (Updated on 17-
XX OCT-2003 to standardise OS field)

SQ Sequence 330 AA;

Query Match 99.8%; Score 1685; DB 2; Length 330;
Best Local Similarity 99.7%; Pred. No. 1.5e-158;

applicant

Matches 329; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVFEKTRPKIAMVSGMTGGTMAFLCSLRELGDVVLFDVVPNMPGKAMDISHNSSVV 60
DB 1 MAVFEKTRPKIAMVSGMTGGTMAFLCSLRELGDVVLFDVVPNMPGKAMDISHNSSVV 60

QY 61 DTGITVYGSNYSYELCKGADVVIITAGITKIPGSKDKEWSRMDLLPVNIKIMRVDGAIAKS 120
DB 61 DTGITVYGSNYSYELCKGADVVIITAGITKIPGSKDKEWSRMDLLPVNIKIMRVDGAIAKS 120

QY 121 YCPNFAVINITNPLDVNVAALQSSGLPHRHCMAAGLQSSRFRMRMIADKLEVSFPRDVQ 180
DB 121 YCPNFAVINITNPLDVNVAALQSSGLPHRHCMAAGLQSSRFRMRMIADKLEVSFPRDVQ 180

QY 181 GMWIGVGHDMVPLSRVATVNGIPLSEFVKKGWIKQEEVDIIVQKTKVAGGEIVRLLGQG 240
DB 181 GMWIGVGHDMVPLSRVATVNGIPLSEFVKKGWIKQEEVDIIVQKTKVAGGEIVRLLGQG 240

QY 241 SAYYAPGASAIQMAESYLKDKRVMVCSYLOGQYGVQNHVYLGVPVCVIGGRGVEKIELE 300
DB 241 SAYYAPGASAIQMAESYLKDKRVMVCSYLOGQYGVQNHVYLGVPVCVIGGRGVEKIELE 300

QY 301 LTAQERQELQSGIDEVKEMOKATAALDASK 330
DB 301 LTAQERQELQSGIDEVKEMOKATAALDASK 330

RESULT 2
ID ADS28805 standard; protein; 320 AA.

AC ADS28805;
DT 02-DEC-2004 (first entry)
DE Bacterial polypeptide #17838.

Recombinant DNA construct; transformed plant; improved plant property;
cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
pathogen tolerance; pest tolerance; plant disease resistance;
cell cycle pathway modification; plant growth regulator;
homologous recombination; seed oil yield; protein yield; carbohydrate;
nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
bacterial polypeptide.

Bacteria.

US2003233675-A1.

18-DEC-2003.

20-FEB-2003; 2003US-00369493.

21-FEB-2002; 2002US-0360039P.

(CAOY/) CAO Y.
(HINK/) HINKLE G J.
(SLAT/) SLATER S C.
(CHEN/) CHEN X.
(GOLD/) GOLDMAN B S.

Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

WPI; 2004-061375/06.

New recombinant DNA construct comprising a promoter positioned to provide
for expression of a polynucleotide encoding a polypeptide from a
microbial source, useful for producing plants with improved properties.

Claim 1; SEQ ID NO 17838; 122pp; English.

The invention relates to a recombinant DNA construct comprising a
promoter functional in a plant cell, where the promoter is positioned to

CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC polynucleotide or polypeptide and growing the transformed plant, where the
CC recombinant DNA construct is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX
SQ Sequence 320 AA;

Query Match 52.2%; Score 880.5; DB 8; Length 320;
Best Local Similarity 53.9%; Pred. No. 1.6e-78;
Matches 173; Conservative 54; Mismatches 89; Indels 5; Gaps 1;

QY 9 RPKIAMVSGMTGGTMAFLCSLRELGDVVLFDVVPNMPGKAMDISHNSSVVDTGITVYG 68

DB 3 RKKIALIGSGMIGGTLAHLAAIKELGDIVLFDIAEGTPOGKALDIAGSGPVEGFDANLKG 62

QY 69 SNSYECLKGADVVIITAGITKIPGSKDKEWSRMDLLPVNIKIMRVEGAAIKYCPNFAVI 128

DB 63 ANSYEDIAGADVCIIVTAGIPRKEG-----MSRDDLLKTNLGVKAVGEGIAAHAPDAFVI 117

QY 129 NITNPLDVNVAALQSSGLPHRHCMAAGLQSSRFRMRMIADKLEVSFPRDVQWIGVGHG 188

DB 118 CITNPLDVNVAALQSSGLPHRHCMAAGLQSSRFRMRMIADKLEVSFPRDVQWIGVGHG 177

QY 189 DHMVPILSRVATVNGIPLSEFVKKGWIKQEEVDIIVQKTKVAGGEIVRLLGQSAIYAPGA 248

DB 178 DTWVPVIEYSTVAGIPVPLIKWGSTQERIDAIIVARTSGSGGEIVALLKTGSAYYAPAT 237

QY 249 SAIQMAESYLKDKRVMVCSYLOGQYGVQNHVYLGVPVCVIGGRGVEKIELELTAQERQE 308

DB 238 SAIAQMAESYLKDKRLLPCCAAHLTGQYGVDDLYVGVPIVIGKDGVERIVEIELNATAKQN 297

QY 309 LOGSDEVKEMOKATAALDAS 329

DB 298 FDSVDVAVKELVAACKSIDAS 318

RESULT 3

ADS22408

ID ADS22408 standard; protein; 319 AA.

XX
AC ADS22408;

XX
DT 02-DEC-2004 (first entry)

XX
DE Bacterial polypeptide #11441.

Recombinant DNA construct; transformed plant; improved plant property;
cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
pathogen tolerance; pest tolerance; plant disease resistance;
cell cycle pathway modification; plant growth regulator;
homologous recombination; seed oil yield; protein yield; carbohydrate;
nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
bacterial polypeptide.

XX
OS Bacteria.

XX PN US2003233675-A1.
 XX PD 18-DEC-2003.
 XX PF 20-FEB-2003; 2003US-00369493.
 XX PR 21-FEB-2002; 2002US-0360039P.
 XX PA (CAOY/) CAO Y.
 XX PA (HINK/) HINKLE G J.
 XX PA (SLAT/) SLATER S C.
 XX PA (CHEN/) CHEN X.
 XX PA (GOLD/) GOLDMAN B S.
 XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX PI WPI; 2004-061375/06.
 XX DR
 XX XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX PS Claim 1; SEQ ID NO 11441; 122pp; English.
 XX XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX SQ .Sequence 319 AA;
 Query Match 51.7%; Score 873.5; DB 8; Length 319;
 Best Local Similarity 53.24; Pred. No. 8.1e-78;
 Matches 167; Conservative 60; Mismatches 82; Indels 5; Gaps 1;
 QY 9 RPKIWMGSMIGTWAFLCSRELGDVLFDDVVPNMPMGKAMDISHNSSVDVTGTVYG 68
 DB 3 RKXIALIGSMIGTGLAHSLKELGDVLFDDIADGIPQKGLDIDIAQSGPVEFNKLSG 62
 QY 69 SNSYECLKGADVVIITAGITKIFPKSKDKEWSRMDLLPVIKIMREYGAATKSCYCPNAFVI 128
 DB 63 ASDYAAIEGADVCIIVTAGVARKFG-----MSRDDLLGLINKVMEQVAGIKKYAPNAFVI 117
 QY 129 NINPLDVMVAALQESGLPHHRCIMAGMDSRFRMTADKLEVPSPRQVQWVGIVGHG 188
 DB 118 CITNPLDAMVWALQKFSGLPKNVGVGNAGVLDARFRLFLAEFNVSDVTFVILGGHG 177
 QY 189 DHMVPLSRVATVNGIPLSEFVKKGIQKEVDIVOKTKVAGSEIVRLILQGSAYYAPGA 248
 DB 178 DTWVPLARYTVGGVPLDVLKRWGLTAERLEQIQIORTDGGAEIVCLLTKTGSAYYAPAA 237
 QY 249 SAIQMAESYLKDRKRVWVCSCYLQGGYGVQNHVGLVPCVIGRGVKEKIILELTAQEROE 308

DB 238 SAIEMAESYLKDKRVLPAHAHLSGQYGVDDMTVGTPTIAGGIERVIELNKSEAA 297
 QY 309 LQGSIDIVEKMQKA 322
 DB 298 FQKSVGAVAGLCEA 311
 RESULT 4
 ADS25456
 ID ADS25456 standard; protein; 320 AA.
 XX AC ADS25456;
 XX DT 02-DEC-2004 (first entry)
 XX DE Bacterial polypeptide #14489.
 XX XX Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX OS Bacteria.
 XX PN US2003233675-A1.
 XX PD 18-DEC-2003.
 XX PF 20-FEB-2003; 2003US-00369493.
 XX PR 21-FEB-2002; 2002US-0360039P.
 XX PA (CAOY/) CAO Y.
 XX PA (HINK/) HINKLE G J.
 XX PA (SLAT/) SLATER S C.
 XX PA (CHEN/) CHEN X.
 XX PA (GOLD/) GOLDMAN B S.
 XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX PI WPI; 2004-061375/06.
 XX DR New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX PS Claim 1; SEQ ID NO 14489; 122pp; English.
 XX XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the

KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX
 OS Bacteria.
 XX
 XX US2003233675-A1.
 XX
 XX 18-DEC-2003.
 XX
 XX 20-FEB-2003; 2003US-00369493.
 XX
 XX 21-FEB-2002; 2002US-0360039P.
 XX
 XX (CAOY/) CAO Y.
 XX (HINK/) HINKLE G J.
 XX (SLAT/) SLATER S C.
 XX (CHEN/) CHEN X.
 XX (GOLD/) GOLDMAN B S.
 XX
 XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 XX
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 XX Claim 1; SEQ ID NO 20748; 122pp; English.
 XX
 XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 XX Sequence 311 AA;
 SQ
 Query Match 51.1%; Score 862.5; DB 8; Length 311;
 Best Local Similarity 53.8%; Pred. No. 9.7e-77;
 Matches 169; Conservative 59; Mismatches 81; Indels 5; Gaps 1;
 QY 9 RPKIANGSGMIGTGAFLCSLRELGDVLFDDVFNPMGKAMDISHNSSVDGTITVG 68
 DB 3 RDKIALIGSGOIGETLAHLVGLKELGDVLFDAEGVPOGKALDIAESSVDGFSKJTG 62
 QY 69 SNSVECLKGADVWITTAGITKIPGSKDEKSRMDLIPVNIKIMREYGAATKSCPNAFVI 128
 DB 63 ANSYEAIEGARVVITVAGVPRKG-----MSRDDLSINLKNWEQVAGIKKYPADFVI 117
 QY 129 NITNPLDVMVAALQESSGLPHHRICGMAGWGLDSSRRFRMTADKLEVSPPRQGMVIGVHG 198

DB 118 CITNPLDMVWALQKASGLPAKKVGMAGVLSARFRYFLADEFNVSVEDTAFVLGGHG 177
 QY 189 DHMVLPSRYATVNGIPLSEFVKKGWIKQBEVDDIVQKTKVAGGEIVRLLGQGSAYAPGA 248
 DB 178 DTMVPLVKYSTVAGIPLDPLVKMGWTSQARLDEIVDRTRNGGAEIVNLLKTGSAYAPAS 237
 QY 249 SAIOAESYLVKDKRVMVCSYLOGQYGVNHYLGVPCVIGGRCVKEKIELELTAEQOE 308
 DB 238 SAIAAESYLVKDKRVMVCSYLOGQYGVNHYLGVPCVIGGRCVKEKIELELTAEQOE 297
 QY 309 LQGSIDEVEMOKA 322
 DB 298 FDRSVAOVGLVEA 311
 RESULT 7
 ADS23027
 ID ADS23027 standard; protein; 322 AA.
 XX
 AC ADS23027;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polypeptide #12060.
 XX
 KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX
 OS Bacteria.
 XX US2003233675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 XX 21-FEB-2002; 2002US-0360039P.
 XX
 XX (CAOY/) CAO Y.
 XX (HINK/) HINKLE G J.
 XX (SLAT/) SLATER S C.
 XX (CHEN/) CHEN X.
 XX (GOLD/) GOLDMAN B S.
 XX
 XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 XX
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 XX Claim 1; SEQ ID NO 12060; 122pp; English.
 XX
 XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,

CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 322 AA;

Query Match 51.1%; Score 862; DB 8; Length 322;
 Best Local Similarity 53.8%; Pred. No. 1.1e-76;
 Matches 174; Conservative 56; Mismatches 87; Indels 8; Gaps 2;
 QY 9 RPKIAMVSGMIGTMAFLCSLRELGDVVLFDVVPNPMGKAMDISHNSSVVDGTGTVYG 68
 DB 3 RNKIALIGSGMIGTMAFLCSLRELGDVVLFDVVPNPMGKAMDISHNSSVVDGTGTVYG 62
 QY 69 SNSYECLKADVVIIITAGITKIPGSKDKWMSRMDLLPVNIIKIMREYGAIAKSYCPNAPVI 128
 DB 63 VNDYAGIEGADVCIIVTAGVPRKPG-----MSRDDLLGINLKVMEQVAGLKKYAPKAFVI 117
 QY 129 NITNPLDVMVAALQESSGLPHHRIICGMAGMLDSSRRFRMIADKLEVPSPDVQGVWIGVG 188
 DB 118 CITNPLDAMVWALQKESGLPKTHVVGAGVLSARFYLAEEFKVSVSDVTAFLVGGHG 177
 QY 189 DHMVPPLSRVATVNGIPLSEFVKVKGWIKQEVDDIVOKTKVAGGEIVRLLGQGSAYYAPGA 248
 DB 178 DSWPMIRISVSGIPLDVLKMGWTSKELDQIVQRTDGGAEIVGLLKTGSAYYAPAA 237
 QY 249 SAIQMAESYLKDKRRVWVCSCYLQGYGVQNHVYLGPCVIGGRGVKEIIIELELTAQERQE 308
 DB 238 SAIQMAESYLKDKRRVLPCCAAHLSGQYGVKGTGVGVVVGAGGVRIIEIDLKSEQKM 297
 QY 309 LQGSIDEVKEMOKA---IAALDASK 330
 DB 298 FESSVATVQGLTEACVKIAPQLASK 322

RESULT 8
 ADS28046
 ID ADS28046 standard; protein; 320 AA.

AC ADS28046;
 XX
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polypeptide #17079.
 XX
 KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.

OS Bacteria.
 XX
 XX US2003233675-A1.
 PN
 XX 18-DEC-2003.
 XX
 XX 20-FEB-2003; 2003US-00369493.
 XX
 XX 21-FEB-2002; 2002US-0360039P.
 PR
 XX (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 XX (GOLD/) GOLDMAN B S.
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.

DR New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 17079; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX

SQ Sequence 320 AA;

Query Match 50.4%; Score 850.5; DB 8; Length 320;
 Best Local Similarity 51.1%; Pred. No. 1.6e-75;
 Matches 164; Conservative 60; Mismatches 92; Indels 5; Gaps 1;
 QY 9 RPKIAMVSGMIGTMAFLCSLRELGDVVLFDVVPNPMGKAMDISHNSSVVDGTGTVYG 68
 DB 3 RAKIALIGAGMIGTMAFLCSLRELGDVVLFDVVPNPMGKAMDISHNSSVVDGTGTVYG 62
 QY 69 SNSYECLKADVVIIITAGITKIPGSKDKWMSRMDLLPVNIIKIMREYGAIAKSYCPNAPVI 128
 DB 63 ANDYADIAGADVCIIVTAGVPRKPG-----MSRDDLLGINLKVMEQVAGLKKYAPKAFVI 117

QY 129 NITNPLDVMVAALQESSGLPHHRIICGMAGMLDSSRRFRMIADKLEVPSPDVQGVWIGVG 188
 DB 118 CITNPLDAMVWALQKESGLPKTHVVGAGVLSARFYLAEEFKVSVSDVTAFLVGGHG 177
 QY 189 DHMVPPLSRVATVNGIPLSEFVKVKGWIKQEVDDIVOKTKVAGGEIVRLLGQGSAYYAPGA 248
 DB 178 DDMVPMVRHSTVGGIPLPELVKQWLSQDKLDAIVERTKRGGEIVALLKTSAGFAPAE 237
 QY 249 SAIQMAESYLKDKRRVWVCSCYLQGYGVQNHVYLGPCVIGGRGVKEIIIELELTAQERQE 308
 DB 238 SAIQMAESYLKDKRRVLPCCAAHLSGQYGVKGTGVGVVVGAGGVRIIEIDLKSEQKM 297
 QY 309 LQGSIDEVKEMOKAIAALDAS 329
 DB 298 FAKSVESVGLMEACKAIDSS 318

RESULT 9
 ADN25360
 ID ADN25360 standard; protein; 320 AA.

XX AC ADN25360;
XX DT 02-DEC-2004 (first entry)
XX DE Bacterial polypeptide #8013.
XX DE
XX KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX OS Bacteria.
XX PN US2003233675-A1.
XX PN 18-DEC-2003.
XX PF 20-FEB-2003; 2003US-00369493.
XX PR 21-FEB-2002; 2002US-0360039P.
XX PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX DR WPI; 2004-061375/06.
XX PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX PS Claim 1; SEQ ID NO 8013; 122pp; English.
XX CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX SQ Sequence 320 AA;
Query Match 50.3%; Score 848.5; DB 8; Length 320;
Best Local Similarity 50.8%; Pred. No. 2.5e-75;
Matches 163; Conservative 64; Mismatches 89; Indels 5; Gaps 1;
9 RPKIAMVGSIGCTTAFGLSLRELGLDVLFDVVPNNPMGKAMDISHSSVWDTGITVYG 68

Db 3 RPKIALIGAGQIGTTLAHLAAIKELGDVLLFDIAEGTPOQKALDIAQSGSPSEGFDAVMKG 62
Qy 69 SNSYECLKGADVVIIITAGITKIPKSKDKWSRMDLLPVIKIMREVGAALKSYCPNAFVI 128
Db 63 ANSYEEIAGADVCIIVTAGVPRKEG-----MSRDDLIIGINLKVMSVGEIGIKAHAPNAFVI 117
Qy 129 NITNPLDVMVAALQESSGLPHHRIICMAGMLDSSRRFRMIADKLEVSPRDVQGMVIGVHG 188
Db 118 CITNPLDMVAALQQFSGLPKAEKVMAGVLDARSFRHFLSVFNVSMDVTAFLVGGHG 177
Qy 189 DHMVPISRYATVNGIPLSEFVKKGWIKOBEVDIVOKTKVAGGEIVRLGQGSAYAPGA 248
Db 178 DTNVPILRYSTVAGIPLPLVQMGWTOEKLDIVORTDGGAEIVGLLTKTGSFAFAPAT 237
Qy 249 SAITMAESYLKDRKRVWVCYLGQGVQVONHYLGPVPCVIGRGVKEIKIELEITAOBRQE 308
Db 238 SALEMAEAYLKQKXLLPCAAVYDGAFLNGVMYGVPTTIAGGIEKIVDKLNDDEQAM 297
Qy 309 LQGSIDEVKEMQKAIADLAS 329
Db 298 FDKSVNAVGLVEACKGIDSS 318
RESULT 10
ADS21490
ID ADS21490 standard; protein; 320 AA.
AC ADS21490;
XX 02-DEC-2004 (first entry)
XX Bacterial polypeptide #10523.
XX KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX OS Bacteria.
XX PN US2003233675-A1.
XX PN 18-DEC-2003.
XX PF 20-FEB-2003; 2003US-00369493.
XX PR 21-FEB-2002; 2002US-0360039P.
XX PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX DR WPI; 2004-061375/06.
XX PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX PS Claim 1; SEQ ID NO 10523; 122pp; English.
XX CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX SQ Sequence 320 AA;

transformed plant having an improved property. The plant is a crop plant such as maize or soybean. The method of producing a transformed plant having an improved property comprises transforming a plant with the recombinant DNA construct and growing the transformed plant, where the polynucleotide or polypeptide is useful for improving plant properties. The recombinant DNA construct is useful for producing plants with improved plant properties, e.g. improved cold, heat or drought tolerance, tolerance to herbicides, extreme osmotic conditions, pathogens or pests, increased resistance to plant disease, better growth rate by modification of the cell cycle pathway with plant growth regulators, increased rate of homologous recombination, modified seed oil or protein yield and/or content, improved yield by modification of carbohydrate, nitrogen or phosphorus use and/or uptake, by modification of photosynthesis or by providing improved plant growth and development under at least one stress condition, improved lignin production or improved galactonannan production. This sequence represents a bacterial polypeptide used in the scope of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 320 AA;

Query Match 50.2%; Score 847.5; DB 8; Length 320;
Best Local Similarity 52.0%; Pred. No. 3.1e-75;
Matches 167; Conservative 61; Mismatches 88; Indels 5; Gaps 1;

QY 9 RPKIAMVSGMIGTMAFLCSLRELGDVVLFDVVPNMPKAMDIHNSVVDGTGITYVG 68
DB 3 RKXIALIGAGNIGGTLAHLAAQKELGDVLFDDVEGVPOGKALDLSQCGPVEGFDANIIG 62
QY 69 SNSYECLKADVVIIITAGITKIPGSKDEWSDMLLPVNIKIMREVGAAIKSCPNAFVI 128
DB 63 TNDYKGIAGADVLIIVTAGVARKG-----MSRDDLLGINLKVNKAVGEGIRDNAPAFVI 117
QY 129 NITNPLDVMVAALQESSGLPHHRIICGVMAGMLDSSRRFRMIADKLEVPDRDQGMVIGVHG 188
DB 118 CITNPLDVMVAALREFSGLPANKVGVAGVLDARSFTFLAWFVGVSIRDVNTFVLGGHG 177
QY 189 DHMVPILSRATVNGIPLSEPVKKGWIKOEVDVIVOKTKVAGGEIVRLILGQGSAYYAPGA 248
DB 178 DTWVPVTVSTVNGIIPVDLVKMGSLTQEKIDAIQVTRSGGGEIVGLLKTGSFAFYAPAA 237
QY 249 SAIQMAESYLKDRKRVWVCSCYLQGGYQVONHYLGVPCVIGRGVEKITELELTAQERQE 308
DB 238 SGTMAEAYLNDKRLIPCAAYVDGEYGVNGLYGVFVLIGANGVEKVEKIEIELDDAKGN 297
QY 309 LQGSIDEVEMQKAIALDAS 329
DB 298 LQVSDAVKLEACKGIDPS 318

RESULT 11
AAR62605
ID AAR62605 standard; protein; 315 AA.
AC AAR62605;
XX AAR62605;

DT 25-MAR-2003 (revised)
DT 18-JUL-1995 (first entry)

XX P. falciparum lactate dehydrogenase.

XX Lactate dehydrogenase; LDH; immunogen; antibody; immunisation; detection;
KW Plasma; serum; malaria.

XX Plasmodium falciparum.

XX WO9424287-A1.

XX 27-OCT-1994.

XX 06-APR-1994; 94WO-US003796.

PR 12-APR-1993; 93US-00046160.
XX (DART-) DARTMOUTH COLLEGE.
XX Bzik DJ, Fox BA;
PI WPI; 1994-341866/42.
XX N-PSDB; AAQ72947.
DR N-PSDB; AAQ72947.
XX Isolated gene encoding lactate dehydrogenase of P. falciparum - and
PT methods for diagnosis and vaccination against malaria.
XX Claim 5; Page 18-19; 35pp; English.
XX This sequence represents P. falciparum lactate dehydrogenase (LDH). The
CC LDH protein and immunogenic fragments of it may be used as an immunogen
CC for antibody generation. The progress of immunisation can be monitored by
CC detection of antibody titres in plasma or serum. Antibodies raised
CC against fragments of LDH can be used to immunise against P. falciparum
CC infection and to detect early malarial infection. See also RR62606-14.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 315 AA;

Query Match 49.1%; Score 828; DB 2; Length 315;
Best Local Similarity 52.6%; Pred. No. 2.7e-73;
Matches 164; Conservative 51; Mismatches 95; Indels 2; Gaps 1;

QY 9 RPKIAMVSGMIGTMAFLCSLRELGDVVLFDVVPNMPKAMDIHNSVVDGTGITYVG 68
DB 4 KAKIVLVSGMIGGVNATLIQVNLGDVVLFDVKNMPHGKALDTSHTNMVAYSNCVSG 63
QY 69 SNSYECLKADVVIIITAGITKIPGSKDEWSDMLLPVNIKIMREVGAAIKSCPNAFVI 128
DB 64 SNTYDLAGADVLIIVTAGITKAPGSKDEWNRDDLPLNNKIMIEIGGHIKNCNPAFII 123
QY 129 NITNPLDVMVAALQESSGLPHHRIICVMAGMLDSSRRFRMIADKLEVPDRDQGMVIGVHG 188
DB 124 VVTNPDVVMVQLLHGHSGVPKNIIGLGGVLDTSRLKYIISQKLVNCPDVAHIVGANG 183
QY 189 DHMVPILSRATVNGIPLSEPVKKGWIKOEVDVIVOKTKVAGGEIVRLILGQGSAYYAPGA 248
DB 184 NKWLLKRYITVGGIPLQEFINNKLISDAELEAIFDRTVNTALEIVNL--HASPYVAPAA 241
QY 249 SAIQMAESYLKDRKRVWVCSCYLQGGYQVONHYLGVPCVIGRGVEKITELELTAQERQE 308
DB 242 AIIEMAESYLKDLKKVLCSTLLEGGYGHSDIFGFTPPVVLGANGVEQVIELQINSEKAK 301
QY 309 LQGSIDEVEMQ 320
DB 302 FDEAIAETKRMK 313

RESULT 12
AAY01690
ID AAY01690 standard; protein; 315 AA.
XX AAY01690;
AC AAY01690;

DT 23-JUN-1999 (first entry)

XX A Plasmodium falciparum lactate dehydrogenase.

XX Lactate dehydrogenase; LDH; antibody; antigenic peptide; malaria;
KW immunization; Plasmodium falciparum infection;
XX Plasmodium falciparum.

OS Plasmodium falciparum.

XX WO9913903-A1.

XX 25-MAR-1999.

XX 08-SEP-1998; 98WO-US018626.

QY 305 EROBLOGSIDEVEMOKAIA 324
DB 293 ERAQLNKSVESVKNVMKVL 312

RESULT 14
ADS28362
ID ADS28362 standard; protein; 314 AA.
AC ADS28362;
DT 02-DEC-2004 (first entry)
XX Bacterial polypeptide #17395.

XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

XX Bacteria.
XX US2003233675-A1.
XX 18-DEC-2003.
XX 20-FEB-2003; 2003US-00369493.
XX 21-FEB-2002; 2002US-0360039P.

PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.

PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 17395; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic

CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 314 AA;

Query Match 42.0%; Score 709; DB 8; Length 314;
Best Local Similarity 44.0%; Pred. No. 1.8e-61;
Matches 140; Conservative 67; Mismatches 97; Indels 14; Gaps 4;

QY 9 RPKIAMVSGMIGGTMAFLCSLRELGDDVVLFDVVPNM---PMGKAMDISHNSSVVDGTG 65
DB 5 RRRKSVIGAGFTGATTALMVAQKELGDDVVLVD-IPQMEGFTKGKALDMLESTPVQGVQDVN 63

QY 66 VYGSNSYECLKGADVVITAGITKIPGKSDKEWSRMDLLPVNIIKIMREYGAIAIKSYCPNA 125
DB 64 ITGTSSEYTKDSDDVVITAGIARKEG-----MSRDDLVSTNAGIMKATVKEVVKHSPNA 118

QY 126 FVINITNPLDVNVAALQESSGLPHHRICGMAGMLSSRRPRMIADKLEVPSPRDVQGMVIG 185
DB 119 YIIILTNPADAMTYTVYKESGPPKRVIGQSGVLDTARFRTFVAQELNLSVEDITGFLVG 178

QY 186 VHGDDHVPILSRVATVNGIPLSEFVKKGWIKQEVDDIVOKTKVAGGEIVRLGCGSAYYA 245
DB 179 GHGDDMVPLIRYSYAGGIPLEKL-----LPQRIIDAIIVERTRGGEIVGLLNGSAYYA 233

QY 246 PGASAIQMAESYLKDKRVMVCYLOGQYQNNHYLGVPCVIGRGVKEIIIELELTAOE 305
DB 234 PAASLAEMVEAILKDKRVLPITAIYLEGEYGEDIVGVPTILGGDGIKFIELDLTDEE 293

QY 306 ROELQGSIDEVEMOKAI 323
DB 294 KATFAKSIESVRNMSAL 311

RESULT 15

AAR94013
ID AAR94013 standard; protein; 329 AA.

AC AAR94013;

DT 16-OCT-2003 (revised)

DT 21-AUG-1996 (first entry)

XX Heat resistant maleate dehydrogenase.

KW Heat resistant maleate dehydrogenase; h-rMAD; NADH; L-aspartic acid;
KW alpha-ketoglutaric acid; glutamine oxalo-transaminase activity; GOT.

OS Geobacillus stearothermophilus; ATCC 12016.

FT Key Location/Qualifiers

FT Misc-difference 185 /note= "Given in the specification as Var"

XX JP08047389-A.

XX 20-FEB-1996.

PF 01-JUL-1994; 94JP-00151045.

XX 02-JUN-1993; 93JP-00164701.

PR 03-JUN-1994; 94JP-00121629.

XX (TOYM) TOYOBO KK.

XX WPI; 1996-166248/17.

DR N-PSDB; AAT17715.

XX Protein having heat resistant malate dehydrogenase activity - and reagent
PT conlg. protein, NADH and L-aspartic and alpha-keto-glutaric acid for
PT determination of glutamine oxalo-transaminase activity.

XX Claim 8; Page 13-15; 17pp; Japanese.

This Page Blank (uspto)